bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The clinical and functional effects of *TERT* variants in myelodysplastic syndrome

Christopher R. Reilly[1,11], Mikko Myllymäki[1,11], Robert Redd[2], Shilpa Padmanaban[3], Druha Karunakaran[1], Valerie Tesmer[3], Frederick D. Tsai[1], Christopher J. Gibson[1], Huma Q. Rana[4], Liang Zhong[5,6], Wael Saber[7], Stephen R. Spellman[8], Zhen-Huan Hu[7], Esther H. Orr[9], Maxine M. Chen[9], Immaculata De Vivo [9,10], Corey Cutler[1], Joseph H. Antin[1], Donna Neuberg[2], Judy E. Garber[4], Jayakrishnan Nandakumar[3], Suneet Agarwal[5,6], R. Coleman Lindsley[1]

[1]Department of Medical Oncology, Division of Hematological Malignancies, Dana-Farber Cancer Institute, Boston, MA, USA
[2]Department of Data Sciences, Dana Farber Cancer Institute, Boston MA, USA
[3]Department of Molecular, Cellular, and Developmental Biology, University of Michigan, Ann Arbor, MI, USA
[4]Division of Population Sciences, Center for Cancer Genetics and Prevention, Dana-Farber Cancer Institute, Boston, MA, USA
[5]Department of Pediatric Oncology, Dana-Farber Cancer Institute, Boston MA, USA
[6]Harvard Stem Cell Institute, Boston MA, USA
[7]Center for International Blood and Marrow Transplant Research, Medical College of Wisconsin, Milwaukee, WI, USA
[8]Center for International Blood and Marrow Transplant Research, National Marrow Donor Program/Be The Match, Minneapolis, MN, USA
[9]Department of Epidemiology, Harvard T.H. Chan School of Public Health, Boston, MA, USA
[10]Channing Division of Network Medicine, Brigham and Women's Hospital and Harvard Medical School, Boston, MA, USA
[11]Contributed equally

Corresponding Author:
R. Coleman Lindsley
Dana-Farber Cancer Institute
450 Brookline Avenue – DA-530C
Boston, MA 02215
E-mail: coleman_lindsley@dfci.harvard.edu

Word count (abstract): 242
Word count (text): 3999
Figures: 5
Tables: 2
References: 64
Running title: *TERT* variants in MDS

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Key points:

- *TERT* rare variants are present in 2.7% of MDS patients and associated with increased non-relapse mortality.
- As a group, *TERT* rare variants have impaired telomere elongation capacity across all structural domains.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Germline pathogenic *TERT* variants are associated with short telomeres and a heightened risk of developing myelodysplastic syndrome (MDS) among patients with dyskeratosis congenita. The prevalence and clinical significance of *TERT* variants in MDS patients undergoing allogeneic stem cell transplant is unknown. We identified *TERT* rare variants in 41 of 1514 MDS patients (2.7%) with genetic and clinical characteristics consistent with a germline origin. *TERT* rare variants occurred in all structural domains and were associated with shorter telomere length ($p<0.001$) and younger age at MDS diagnosis ($p=0.04$). No patients with a *TERT* rare variant had a known telomere biology disorder. In multivariable analyses, *TERT* rare variants were associated with inferior overall survival ($p=0.034$) driven by an increased incidence of non-relapse mortality (NRM) ($p=0.015$). Deaths from non-infectious pulmonary causes were more frequent in patients with a *TERT* rare variant. Across all major structural domains, *TERT* rare variants demonstrated impaired capacity to elongate telomeres in a cell-based assay. Using a homology model of human TERT bound to the shelterin protein TPP1, we inferred that *TERT* rare variants likely disrupt domain-specific functions. Our results indicate that the contribution of *TERT* rare variants to MDS pathogenesis and NRM risk is underrecognized. Systematic screening for *TERT* rare variants in MDS patients regardless of age or clinical suspicion could identify clinically inapparent telomere biology disorders and improve transplant outcomes through risk-adapted approaches. Cell-based functional characterization of *TERT* rare variants may facilitate *TERT*-specific variant classifications guidelines with broad clinical applicability.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Impaired telomere maintenance is implicated in the pathogenesis of myelodysplastic syndrome (MDS),[1–5] for which allogeneic hematopoietic stem cell transplantation (HSCT) is the only potential cure.[6,7] Shorter pre-transplant blood telomere length is independently associated with an increased risk of early non-relapse mortality (NRM) in MDS patients,[8] but the genetic determinants of telomere length in MDS are incompletely characterized.

Germline pathogenic variants affecting telomerase- and telomere-associated proteins cause a global impairment in telomere maintenance and short telomeres in all tissues.[1,9–11] Individuals with dyskeratosis congenita (DC), an early-onset syndromic presentation of a telomere biology disorder, have characteristic mucocutaneous features, bone marrow failure,[2,12] and a markedly increased risk of developing MDS or acute myeloid leukemia.[2,4,13,14] In contrast, adult patients with a telomere biology disorder more frequently present with aplastic anemia,[15–17] idiopathic pulmonary fibrosis,[18–23] and liver cirrhosis.[21,24,25] Affected families may show anticipation, marked by changes in the onset, phenotype, and severity of clinical disease across successive generations.[21,26–28] Clinical suspicion for a telomere biology disorder is based on the presence of syndromic features and disease phenotypes in relatives.[3,22] However, clinical manifestations are highly variable, and up to 40% of affected patients lack a family history of hematologic, pulmonary, or hepatic abnormalities.[4] In contrast to DC, the risk of developing MDS in older adults with late-presenting or unrecognized telomere biology disorders is unknown.

TERT is the most frequently mutated gene among patients with a telomere biology disorder[4] and can cause disease in an autosomal dominant form.[1] *TERT* encodes telomerase reverse transcriptase that binds to the telomerase RNA component (TERC) and functions within the multi-subunit telomerase holoenzyme complex to extend telomere ends during DNA replication.[9,29] Telomerase is composed of four structural domains with distinct functional

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

roles:[30,31] the telomerase essential/N-terminal (TEN) domain, telomerase RNA-binding domain (TRBD), reverse transcriptase domain (RTD), and C-terminal extension (CTE) domain. Disease-associated germline *TERT* variants are predominantly missense substitutions and occur within all structural domains.[32] Novel *TERT* missense variants are classified as variants of unknown significance (VUS) by consensus guidelines in the absence of additional supporting computational or functional evidence of pathogenicity.[33,34] However, *in silico* prediction algorithms have limited utility in assessing genotype-phenotype relationships for missense substitutions.[35–37] Furthermore, the cellular effect of *TERT* variants may not be reflected by *in vitro* functional assays.[38–40]

The prevalence and clinical significance of *TERT* variants among MDS patients unselected for suspicion of a telomere biology disorder are unknown. Here we analyzed the clinical and functional effects of *TERT* variants in a registry-level cohort of MDS patients who underwent allogeneic HSCT.

Methods

Patients and samples

We previously described a cohort of 1514 MDS patients who were enrolled in The Center for International Blood and Marrow Transplant Research repository and research database who had banked whole peripheral blood DNA samples.[41] The median follow-up time for censored patients was 5.0 years. A separate cohort of 401 adult patients with non-Hodgkin lymphoma (NHL) patients treated with high-dose chemotherapy with autologous stem cell rescue at the Dana–Farber Cancer Institute had banked mobilized whole peripheral blood DNA samples.[41] This study was conducted with the approval of the institutional review board at the Dana–Farber Cancer Institute.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

DNA library preparation, sequencing, and variant annotation

In the MDS cohort, we sequenced the *TERT* coding region (exons 1-16) and known germline single nucleotide polymorphisms (SNPs). In the NHL cohort, we sequenced *TERT* in DNA extracted from mobilized whole peripheral blood samples. The genetic analysis was completed and locked prior to merging with clinical data. Detailed sequencing methods are in the supplementary methods.

Telomere Length Measurements

Relative telomere length of MDS patients was measured by quantitative polymerase chain reaction (qPCR) as previously reported.[8] K562 cell DNA was extracted using the QIAamp Blood Mini kit (Qiagen). Telomere length was measured by two orthogonal methods: 1) qPCR[8] and 2) telomere restriction fragment (TRF) analysis using TeloTAGGG Telomere Length Assay (Sigma Aldrich) according to the manufacturer's protocol.

Plasmids, cloning, site-directed mutagenesis

Human TERT cDNA[39] was cloned into the gateway pDONR221 plasmid (Invitrogen). *TERT* variants were generated using the Q5 Site-Directed Mutagenesis Kit (NEB) and primers designed using NEBase changer (Table S1). Transfer of each construct to the pCW57.1 destination plasmid (Addgene) was performed using LR clonase (ThermoFisher) and the complete sequence was confirmed by Sanger sequencing.

Lentivirus production and cell line generation

Lentivirus for each *TERT* variant was produced in HEK 293T cells. *TP53*-repaired K562 cells (gift from the Ebert lab)[42] were transduced at MOI ~ 1 followed by puromycin selection (2ug/mL) to generate bulk cell lines. K562 cells were treated with doxycycline (1ug/mL) for 27 days. RPMI media was added every 2 days and cells were split every 3 days.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Western blotting for hTERT expression

Protein extracts were prepared in Laemmli 2X buffer and western blots performed using SDS gels (BioRad) and Trans-Blot Turbo Transfer System. hTERT expression was visualized using hTERT antibody (Rockland; Cat# 600-401-252; 1:1,000 dilution) with anti-rabbit secondary antibody. B-actin was used as loading control (Abcam; Cat# ab20272; 1:10,000 dilution). Chemiluminescence images were obtained using Bio-Rad ChemiDoc and analyzed using Image Lab software.

Structural modeling of residues mutated in TERT rare variants

A homology model of human telomerase bound to TPP1-OB and parts of TERC was generated using the cryo-EM structure of the *Tetrahymena thermophilia* telomerase holoezyme[43,44], the crystal structure of human TPP1-OB[45] (as described in Tesmer and Smith et al, *PNAS* 2019), as well as additional crystal structures of TERT/TERC domains from various species[43,45–47] using Phyre 2.[48] *TERT* rare variant positions were manually annotated in the final homology model. Details of the homology model are listed in supplemental methods.

Statistical analyses

Fisher's exact test was used to test for association between pairs of categorical variables. The Wilcoxon rank-sum test was used to assess a location shift in the distribution of continuous variables between two groups. For associations with ordered categorical variables, Cochran-Armitage trend test was used for singly-ordered contingency tables. All p-values were two-sided.

Overall survival was defined as the time from transplant until death from any cause. Subjects not confirmed dead were censored at the time last known to be alive. Differences in survival

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

curves were assessed using log-rank tests. NRM was defined as death without relapse. NRM, with relapse as a competing risk, was assessed with the use of Gray's test. For relapse, death without relapse was considered a competing risk. Univariate and multivariable analyses of OS were performed using Cox regression. OS estimates were calculated using the method of Kaplan-Meier and reported with 95% CIs based on Greenwood's formula. Hazard ratios with 95% CIs and Wald P values were reported for covariates in multivariable Cox models. Multivariable models for competing risks of relapse and NRM were generated using the Fine and Gray method.

Results

TERT variants in MDS and NHL cohorts

In total, we identified 270 nonsynonymous *TERT* coding variants among the MDS and NHL cohorts (Figure 1A). Pathogenic genetic variants are observed infrequently in the general population due to strong negative selection.[34,49] Therefore, we grouped *TERT* variants based on their maximum gnomAD population allele frequency in any reference population,[50] where "common" variants had a maximum allele frequency ≥0.001 and "rare" variants had a maximum allele frequency <0.001. Using this approach, 228 variants (84.4%) were classified as common and 42 variants (R1086H occurred in two MDS patients) as rare (15.6%) (Figure 1A). The frequency of *TERT* common variants was similar in the MDS and NHL cohorts (11.9% vs. 12.0%, p=0.79), and primarily included the SNPs p.A279T (rs61748181), p.H412Y (rs34094720), p.E441del (rs377639087), and p.A1062T (rs35719940) (Figure 1C and Figure S1). In contrast, *TERT* rare variants were significantly more common in patients with MDS (2.7% MDS vs. 0.25% NHL, p <0.001)(Figure 1B).

In MDS patients, *TERT* rare variants occurred within all structural domains: RTD (n=15), CTE domain (n=10), TRBD (n=8), TEN domain (n=4), and linker region between TEN and TRBD

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(n=3) (Figure 1C). The majority of variants were missense substitutions (39 of 40) with one splice site variant (c.1770-2A>G) (Table 1). According to ACMG/AMP[33] and Sherloc criteria[34], 39 variants were classified as variants of uncertain significance (VUS) and R865C was classified as likely pathogenic. Twenty-three variants (57.5%) were absent from gnomAD, 17 variants were listed in ClinVar (42.5%), and 11 variants (28.2%) have been reported in patients with telomere biology disorders (Table S2). *In silico* Combined Annotation-Dependent Depletion (CADD)[49] PHRED-like scores > 20 and ClinPred[51] scores >0.5 were observed in 54% and 51% of variants, respectively (Tables 1 and S2). In contrast, *TERT* common variants were classified as benign or likely benign (Table S3), and most have been experimentally determined to have comparable activity to wild type TERT.[40]

TERT rare variants and clinical characteristics

Patients with a *TERT* rare variant had shorter telomere length (ddCT=0.405 vs. 0.507, p<0.001) and were younger at MDS diagnosis (median age 52 vs. 59, p=0.03) than those without a *TERT* variant (Figure 2A and B). All other clinical characteristics, including Karnofsky performance status, hematopoietic stem cell transplant comorbidity index (HCT-CI) score, peripheral blood counts at transplant, frequency of therapy-related MDS, and IPSS-R risk group were similar in patients with and without a *TERT* rare variant (Tables 1 and S6). Patients with *TERT* common variant were similar to those without any *TERT* variant with respect to telomere length (ddCT=0.509 vs. 0.507, p=0.80), age at MDS diagnosis (median age 59 vs. 59, p=0.72), and other clinical characteristics (Figure 2A and B, Table S6).

In the absence of available constitutional reference tissue, we used genetic characteristics to determine whether *TERT* rare variants were likely present in the germline or likely acquired somatically within the malignant clone. Germline variants are present in all cells and have a variant allele fraction (VAF) around 0.5 (heterozygous) or 1 (homozygous), while somatic

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

mutations are present only in a subset of clonal cells with a wider range of VAF that falls below 0.5 in diploid cells. *TERT* rare variants and control germline SNPs had a median VAF of 0.48 (95% CI 0.457-0.502) and 0.56 (95% CI 0.557-0.564), respectively (Figure 2C and Table S2), whereas MDS somatic mutations typically present in the founding clone displayed lower median VAFs (*DNMT3A*: 0.10, *TET2*: 0.16, *SRSF2*: 0.25, *U2AF1*: 0.20, *ASXL1*: 0.18). Additionally, the VAF of *TERT* rare variants did not vary with the proportion of blood lymphocytes (Figure 2D), indicating that the variants were present in all nucleated cells, including both the clonal myeloid compartment and the typically non-clonal lymphoid compartment.

TERT rare variants and clinical outcomes

To determine whether *TERT* variant status was associated with clinical outcomes after transplantation, we evaluated overall survival and the cumulative incidences of relapse and NRM. Among 41 MDS patients with a *TERT* rare variant, overall survival (OS) was 24% at 5 years (Figure 3A), and the cumulative incidences of NRM and relapse at 5 years were 52.5% and 27.5%, respectively (Figures 3B and C). In multivariable analysis, the presence of a *TERT* rare variant compared with the absence of a *TERT* rare variant was associated with inferior overall survival (HR for death 1.50, 95% CI 1.04-2.20, p=0.03) and an increased rate of NRM (HR 1.75, 95% CI 1.13-2.72, p=0.01) but not a higher rate of relapse (HR 0.78, 95% CI 0.42-1.16, p=0.44) (Figure 3D). Non-genetic factors also impacted the rate of non-relapse mortality in this model, including recipient age (per 10-year increase, HR 1.23, 95% CI 1.14-1.33, p<0.01) and Karnofsky performance score <90 (1.23, 1.02-1.53, p=0.03). The results of the multivariable Cox model for overall survival and the Fine–Gray model for the rates of relapse and NRM, along with adjusted covariates for each model, are provided in Table S8.

Primary disease (32%), non-infectious pulmonary causes (21%), and infections (18%) were the most common causes of death in patients with a *TERT* rare variant (Table S7). Among these,

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

non-infectious pulmonary causes of death occurred more frequently in patients with a *TERT* rare variant compared to those without a *TERT* rare variant. Five of six patients with a *TERT* rare variant and non-infectious pulmonary cause of death received myeloablative conditioning. The incidence of pre-transplant pulmonary dysfunction, as defined by HCT-CI criteria[52], was not significantly different between patients with (50%) and without a *TERT* rare variant (39%) (p=0.27).

Functional effects of TERT rare variants

We next determined the impact of *TERT* rare variants on telomere elongation in human cells (Figure 4A). Doxycycline-inducible expression of wild type TERT in K562 AML cells resulted in progressive increase in telomere length over 27 days, while expression of luciferase or a known catalytically impaired *TERT* variant ($TERT^{V694M}$)[15,40] resulted in minimal change in telomere length (Figure 4B). Bulk cell lines were generated for all 39 *TERT* missense rare variants and one *TERT* common variant (p.A279T). TERT expression was consistent throughout the experiment (Figure 4A and S6). Telomere elongation capacity was calculated as the change in qPCR telomere length from day 0 to day 27 normalized to that of wild type TERT (Figure 4C and S7).

Most TERT rare variants exhibited impaired capacity to elongate telomeres compared with wild type TERT (Figure 4C; Figures S3 and S5). Eighteen variants (46.2%) displayed severely impaired telomere elongation capacity (<25% of wild type), including 10 of 11 previously reported variants associated with telomere biology disorders (Table S2). Intermediate telomere elongation capacity (25-75% of wild type) was observed for seventeen variants (43.6%). C76S, G135R, G306S, and R1086H exhibited preserved telomere elongation capacity (>75% of wild type). Severely impaired variants occurred within all major structural domains of TERT: TEN (2

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of 4), TRBD (3 of 8), RTD (9 of 15), CTE (4 of 10). As a group, linker variants had a modest impact on telomere elongation capacity (H263H: 53%, H296P: 71%, G306S: 82%).

Structural analysis of TERT rare variants

To study the potential structural effects of *TERT* rare variants, we constructed a homology model of human telomerase bound to the OB domain of the shelterin protein TPP1 as described in methods and supplement. In this model, the TRBD, RTD, and CTE domains form a closed ring structure that is consistent with low-resolution cryo-EM data from human telomerase (Figure 5A).[53] The TEN domain straddles the insertion in fingers domain (IFD) portion of the RTD and contacts the CTE domain, thereby trapping the TERC template and DNA substrate within the active site (TERC and DNA omitted in Figure 5A for simplicity). The TPP1 OB domain docks at the TEN-IFD interface of TERT, as described previously.[53]

Of the fifteen *TERT* rare variants that map to the RTD, ten localize within the catalytic core region (Figure 5B) that is structurally similar to other known reverse transcriptases. [31] Four of these ten variants (R622H, G715D, R865C, and V867M) are proximal to the active site pocket and the RNA-DNA duplex (*left* panel of Figure 5B; see Figure 5C for R662). In contrast, variants S663G, R669W, R698Q (not shown), G847S, D848N, and T917M map distal to the active site (*center* panel of Figure 5B; see *left* panel for T917). We also identified five RTD variants that lie within the IFD.[53] The IFD consists of two bracing helices with an intervening TERT-specific "TRAP" subdomain that entraps the template-primer duplex within the active site (center panel of Figure 5B). Q722R resides at the base of the N-terminal bracing helix, and four variants (V777M, P771L, V741L and L766S) lie within the TRAP region. Notably, V777 resides within an α helix that contacts the TEN domain, while P771 lies immediately adjacent to it.

The TEN domain facilitates telomere repeat addition processivity (RAP) of telomerase and mediates telomerase recruitment to the ends of chromosomes through specific interactions with

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the TERT IFD-TRAP and the N-terminal OB domain of TPP1.[53] Four variants (C76S, V84M, G110A (not modeled), and G135R) localize to the TEN domain and are proximal to a region implicated in recognizing TPP1 (Figure 5D).

The CTE domain, known as the thumb domain in other polymerases, is composed of four highly conserved motifs (E-I, E-II, E-III, E-IV) essential for biological activity through RNA-DNA duplex binding as well as RAP.[54] Eight of ten CTE variants localize to E-I (R951W, S984R, L994F, A1014P), E-II (S1041F), and E-III (R1086H, R1086C, V1090M) (Figure 5E). At opposite ends of a loop connecting E-III and E-IV, R1105W sits in close proximity to TERC, whereas T1110M resides on a solvent-exposed face.

The TRBD interacts extensively with TERC via the CR4/5 domain and pseudoknot/template region.[46] Among the eight TRBD variants, R485C, E484K, and A532T (not shown) localize to alpha-helices in contact with the CR4/5 domain of TERC in the homology model, whereas T567M and K570R reside on a hydrophilic loop in close proximity to the RNA-DNA duplex (Figure 5C). V461E targets a residue buried within the protein hydrophobic core. In contrast, variant V435L (not shown) occurs at a poorly conserved residue within an unstructured region. The remaining three variants (R263H, H296P, and G306S) localize to a weakly conserved linker region that connects the TEN domain to the TRBD (not modeled).

Discussion

The prevalence, prognostic significance, and functional effects of *TERT* variants in MDS patients have not been systematically evaluated. In this study, we identified all *TERT* variants in a registry-level cohort of 1514 MDS patients unselected for suspicion of telomere biology disorder and studied their clinical and functional consequences. *TERT* variants that are frequent in population databases, such as A279T, H412Y, E441del, and A1062T, had no apparent phenotypic consequences, consistent with studies showing that common polymorphisms do not

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

contribute measurably to telomere-related diseases.[40] In contrast, *TERT* rare variants (<0.1% in all reference populations) were present in 2.7% of MDS patients and were associated with characteristics of disease-causing germline mutations, including shorter telomere length, younger age at MDS diagnosis, and impaired telomere elongation capacity in human cells.

As a group, patients with a *TERT* rare variant had poor survival after allogeneic HSCT owing to an increased risk of NRM. The prognostic impact of *TERT* rare variants was independent of established clinical predictors of NRM, such as recipient age, Karnofsky performance status, HCT-CI score, donor-recipient HLA matching, and conditioning intensity. In particular, patients with a *TERT* rare variant were more likely to die from a non-infectious pulmonary cause than those without a *TERT* rare variant. This increased risk of NRM and post-transplantation pulmonary complications evokes studies that have reported high rates of NRM and fatal post-transplant pulmonary complications in patients with dyskeratosis congenita.[2,55–58] In this context, a global defect in telomere maintenance and constitutionally short telomeres may render patients susceptible to non-hematopoietic end-organ toxicity after conditioning with radiation or DNA alkylating agents.

Most germline *TERT* variants observed in telomere biology disorder patients are missense substitutions.[32] In the absence of compelling family history or functional data, novel variants thus present a clinical dilemma, where accurate variant classification relies on multi-tiered evidence to support interpretation.[33,34,59] Most of the *TERT* rare variants we identified in this MDS cohort were classified as VUS and would not be definitively actionable in clinical practice. *In silico* modeling approaches have shown limited utility in predicting the effects of missense variants in patients.[35] Moreover, *in vitro* functional assays may fail to reveal non-enzymatic defects that are essential for *in vivo* activity, including nuclear localization,[60] holoenzyme assembly,[61] and telomerase recruitment to telomeres.[53,62] We therefore determined the functional effects of all

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

candidate *TERT* rare variants in a cell-based assay with *in vivo* telomere extension as the read-out. Using this approach, we showed that 90% of rare variants caused a quantifiable defect in telomere elongation compared to wild type TERT, whereas the SNP A279T had preserved function. The degree of impairment among rare missense substitutions was variable, with 18 having severe functional effect (<25% of wild type telomere elongation capacity) and 17 having intermediate function (25-75% of wild type). Variants that displayed preserved telomere elongation capacity in our assay (>75% of wild-type) may represent private genetic variants with no significant biological impact. Conversely, a mild functional impairment of these variants may not be evident in our assay, but nevertheless contribute to clinical disease due to genetic anticipation or in combination with other factors increasing hematopoietic cell turnover.

TERT rare variants were distributed across multiple domains, suggesting that there are multiple mechanisms by which TERT variants can impair telomere extension. By correlating the functional data with the human telomerase homology model, we inferred the mechanistic basis of each variant's effect. For example, while all 15 *TERT* rare variants within the RTD demonstrated reduced telomere extension, 10 were in proximity to the active site and potentially impair catalysis directly,[15] while the 5 IFD variants likely alter TPP1-dependent telomere association.[63] In this regard, the two variants within the TEN domain that showed severely reduced telomere elongation capacity (V84M and G110A) are also positioned to likely disrupt TPP1-mediated telomerase recruitment.[53] In contrast, variants within the TRBD and CTE likely impair interactions with regions of TERC, including its CR4/5 and template/pseudoknot domains, that are important for both ribonucleoprotein complex stability and catalysis.[30] Our complementary functional and structural analysis provides a powerful framework for evaluation of novel *TERT* variants and may enhance establishment of TERT-specific variant classification guidelines with broad clinical applicability.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Together, our results indicate that *TERT* rare variants identify a group of MDS patients who may have an unrecognized telomere biology disorder. This conclusion is based on functional characterization of all candidate variants, telomere length measurements in primary patient samples, and annotation of clinical characteristics including age, comorbidities, and toxicity outcomes in a registry-level cohort. Our analysis is limited by the unavailability of germline reference tissue and absence of detailed family history and clinical examination. Importantly, no patient with a *TERT* rare variant had a clinical diagnosis of dyskeratosis congenita, and pre-transplant clinical characteristics such as pulmonary and hepatic function, peripheral blood counts, and history of aplastic anemia were similar among patients with or without a *TERT* rare variant. This observation is consistent with previous reports that adults with telomere biology disorder rarely exhibit syndromic features and affected patients often lack a family history of hematologic, pulmonary, or hepatic abnormalities.[4,21] Indeed, MDS has been reported to be a late presenting disease manifestation in patients with telomere biology disorders.[2,4,21]

Our results are consistent with the varied clinical presentation of patients with a telomere biology disorder and indicate that clinical criteria alone may be inadequate to identify all patients with germline *TERT* mutations. Notably, 90% of MDS patients with *TERT* rare variants were adults older than 40 years of age and there appeared to be no upper age limit. Further, the predictive value of telomere length thresholds in identifying patients with a *TERT* mutation has been shown to be poor in older patients, where the telomere length of affected patients overlaps with the lower range of the normal aging control population.[21] The unexpectedly high prevalence of unrecognized and clinically significant telomerase alterations among adult MDS patients thus raises the possibility that routine *TERT* sequencing should be incorporated into standard MDS diagnostic evaluation irrespective of age, clinical presentation, or family history. The results of screening could directly inform transplant donor selection by enabling exclusion of candidate related donors who share the germline allele. Further, pre-transplantation referral for evaluation

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

of comorbid pulmonary or hepatic disease and mandating less intensive conditioning regimens could mitigate the elevated risk of NRM. Such a strategy would require multidisciplinary assessment and gene-specific guidelines for variant classification to guide clinical decision-making.[64]

The frequency of *TERT* rare variants does not fully account for the adverse effect of short telomere length on non-relapse mortality in adult MDS patients.[8] Genetic alterations to other components of the telomerase and shelterin complexes are also associated short telomeres and clinical disease.[1,2,32] Unbiased sequencing of these genes, paired with telomere length measurement and clinical outcomes, may reveal additional gene variants associated with MDS predisposition and similarly inferior transplant outcomes.

In conclusion, we show that *TERT* rare variants impair telomere elongation in cells and are associated with shorter telomeres, younger age at diagnosis, and an increased risk of NRM in MDS patients undergoing allogeneic transplantation. These results suggest that unrecognized telomere biology disorders contribute to the pathogenesis of MDS. Identifying *TERT* variants via systematic genetic screening in MDS patients of all ages and regardless of clinical suspicion could impact clinical care by informing donor selection, family counseling, and mitigation of NRM risk.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

References

1. Savage SA, Bertuch AA. The genetics and clinical manifestations of telomere biology disorders. *Genet. Med.* 2010;12(12):753–764.
2. Agarwal S. Evaluation and Management of Hematopoietic Failure in Dyskeratosis Congenita. *Hematol. Oncol. Clin. North Am.* 2018;32(4):669–685.
3. Schratz KE, DeZern AE. Genetic Predisposition to Myelodysplastic Syndrome in Clinical Practice. *Hematol. Oncol. Clin. North Am.* 2020;34(2):333–356.
4. Schratz KE, Haley L, Danoff SK, et al. Cancer spectrum and outcomes in the Mendelian short telomere syndromes. *Blood*. 2020;135(22):1946–1956.
5. Kirwan M, Vulliamy T, Marrone A, et al. Defining the pathogenic role of telomerase mutations in myelodysplastic syndrome and acute myeloid leukemia. *Hum. Mutat.* 2009;30(11):1567–1573.
6. de Witte T, Bowen D, Robin M, et al. Allogeneic hematopoietic stem cell transplantation for MDS and CMML: recommendations from an international expert panel. *Blood*. 2017;129(13):1753–1762.
7. Bartenstein M, Deeg HJ. Hematopoietic stem cell transplantation for MDS. *Hematol. Oncol. Clin. North Am.* 2010;24(2):407–422.
8. Myllymäki M, Redd RA, Reilly CR, et al. Short Telomere Length Predicts Non-Relapse Mortality after Stem Cell Transplantation for Myelodysplastic Syndrome. *Blood*. 2020;
9. Blackburn EH, Epel ES, Lin J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. *Science*. 2015;350(6265):1193–1198.
10. Demanelis K, Jasmine F, Chen LS, Chernoff M, Tong L. Determinants of telomere length across human tissues. *bioRxiv*. 2019;
11. Grill S, Nandakumar J. Molecular mechanisms of telomere biology disorders. *J. Biol. Chem.* 2020;296:100064.
12. Savage SA. Dyskeratosis Congenita. 1993;
13. Alter BP, Giri N, Savage SA, Rosenberg PS. Cancer in the National Cancer Institute inherited bone marrow failure syndrome cohort after fifteen years of follow-up. *Haematologica*. 2018;103(1):30–39.
14. Marsh JCW, Gutierrez-Rodrigues F, Cooper J, et al. Heterozygous RTEL1 variants in bone marrow failure and myeloid neoplasms. *Blood Adv*. 2018;2(1):36–48.
15. Yamaguchi H, Calado RT, Ly H, et al. Mutations in TERT, the gene for telomerase reverse transcriptase, in aplastic anemia. *N. Engl. J. Med.* 2005;352(14):1413–1424.
16. Calado RT, Regal JA, Kajigaya S, Young NS. Erosion of telomeric single-stranded overhang in patients with aplastic anaemia carrying telomerase complex mutations. *Eur. J. Clin. Invest.* 2009;39(11):1025–1032.
17. Collopy LC, Walne AJ, Vulliamy TJ, Dokal IS. Targeted resequencing of 52 bone marrow failure genes in patients with aplastic anemia reveals an increased frequency of novel variants of unknown significance only in SLX4. *Haematologica*. 2014;99(7):e109–11.
18. Tsakiri KD, Cronkhite JT, Kuan PJ, et al. Adult-onset pulmonary fibrosis caused by mutations in telomerase. *Proc. Natl. Acad. Sci. U. S. A.* 2007;104(18):7552–7557.
19. Armanios MY, Chen JJ-L, Cogan JD, et al. Telomerase mutations in families with idiopathic pulmonary fibrosis. *N. Engl. J. Med.* 2007;356(13):1317–1326.
20. Newton CA, Batra K, Torrealba J, et al. Telomere-related lung fibrosis is diagnostically heterogeneous but uniformly progressive. *Eur. Respir. J.* 2016;48(6):1710–1720.
21. Alder JK, Hanumanthu VS, Strong MA, et al. Diagnostic utility of telomere length testing in a hospital-based setting. *Proc. Natl. Acad. Sci. U. S. A.* 2018;115(10):E2358–E2365.
22. Diaz de Leon A, Cronkhite JT, Katzenstein A-LA, et al. Telomere lengths, pulmonary fibrosis and telomerase (TERT) mutations. *PLoS One*. 2010;5(5):e10680.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

23. Petrovski S, Todd JL, Durheim MT, et al. An Exome Sequencing Study to Assess the Role of Rare Genetic Variation in Pulmonary Fibrosis. *Am. J. Respir. Crit. Care Med.* 2017;196(1):82–93.
24. Gorgy AI, Jonassaint NL, Stanley SE, et al. Hepatopulmonary syndrome is a frequent cause of dyspnea in the short telomere disorders. *Chest*. 2015;148(4):1019–1026.
25. Donaires FS, Scatena NF, Alves-Paiva RM, et al. Telomere biology and telomerase mutations in cirrhotic patients with hepatocellular carcinoma. *PLoS One*. 2017;12(8):e0183287.
26. Parry EM, Alder JK, Qi X, -L. Chen JJ, Armanios M. Syndrome complex of bone marrow failure and pulmonary fibrosis predicts germline defects in telomerase. *Blood*. 2011;117(21):5607–5611.
27. Armanios M, Chen J-L, Chang Y-PC, et al. Haploinsufficiency of telomerase reverse transcriptase leads to anticipation in autosomal dominant dyskeratosis congenita. *Proc. Natl. Acad. Sci. U. S. A.* 2005;102(44):15960–15964.
28. Vulliamy T, Marrone A, Szydlo R, et al. Disease anticipation is associated with progressive telomere shortening in families with dyskeratosis congenita due to mutations in TERC. *Nat. Genet.* 2004;36(5):447–449.
29. Greider CW, Blackburn EH. Telomeres, telomerase and cancer. *Sci. Am.* 1996;274(2):92–97.
30. Nguyen THD, Tam J, Wu RA, et al. Cryo-EM structure of substrate-bound human telomerase holoenzyme. *Nature*. 2018;557(7704):190–195.
31. Smith EM, Pendlebury DF, Nandakumar J. Structural biology of telomeres and telomerase. *Cell. Mol. Life Sci.* 2020;77(1):61–79.
32. Podlevsky JD, Bley CJ, Omana RV, Qi X, Chen JJ-L. The telomerase database. *Nucleic Acids Res.* 2008;36(Database issue):D339–43.
33. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet. Med.* 2015;17(5):405–423.
34. Nykamp K, Anderson M, Powers M, et al. Sherloc: a comprehensive refinement of the ACMG–AMP variant classification criteria. *Genet. Med.* 2017;19(10):1105–1117.
35. Miosge LA, Field MA, Sontani Y, et al. Comparison of predicted and actual consequences of missense mutations. *Proc. Natl. Acad. Sci. U. S. A.* 2015;112(37):E5189–98.
36. Ernst C, Hahnen E, Engel C, et al. Performance of in silico prediction tools for the classification of rare BRCA1/2 missense variants in clinical diagnostics. *BMC Med. Genomics*. 2018;11(1):35.
37. Ghosh R, Oak N, Plon SE. Evaluation of in silico algorithms for use with ACMG/AMP clinical variant interpretation guidelines. *Genome Biol.* 2017;18(1):225.
38. Armbruster BN, Banik SS, Guo C, Smith AC, Counter CM. N-terminal domains of the human telomerase catalytic subunit required for enzyme activity in vivo. *Mol. Cell. Biol.* 2001;21(22):7775–7786.
39. Counter CM, Hahn WC, Wei W, et al. Dissociation among in vitro telomerase activity, telomere maintenance, and cellular immortalization. *Proc. Natl. Acad. Sci. U. S. A.* 1998;95(25):14723–14728.
40. Zaug AJ, Crary SM, Fioravanti MJ, Campbell K, Cech TR. Many disease-associated variants of hTERT retain high telomerase enzymatic activity. *Nucleic Acids Research*. 2013;41(19):8969–8978.
41. Gibson CJ, Lindsley RC, Tchekmedyian V, et al. Clonal hematopoiesis associated with adverse outcomes following autologous stem cell transplantation for non-Hodgkin lymphoma. 2016;
42. Boettcher S, Miller PG, Sharma R, et al. A dominant-negative effect drives selection of

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

TP53 missense mutations in myeloid malignancies. *Science*. 2019;365(6453):599–604.
43. Jiang J, Wang Y, Sušac L, et al. Structure of Telomerase with Telomeric DNA. *Cell*. 2018;173(5):1179–1190.e13.
44. Jiang J, Chan H, Cash DD, et al. Structure of Tetrahymena telomerase reveals previously unknown subunits, functions, and interactions. *Science*. 2015;350(6260):aab4070.
45. Wang F, Podell ER, Zaug AJ, et al. The POT1–TPP1 telomere complex is a telomerase processivity factor. *Nature*. 2007;445(7127):506–510.
46. Huang J, Brown AF, Wu J, et al. Structural basis for protein-RNA recognition in telomerase. *Nat. Struct. Mol. Biol.* 2014;21(6):507–512.
47. Hoffman H, Rice C, Skordalakes E. Structural Analysis Reveals the Deleterious Effects of Telomerase Mutations in Bone Marrow Failure Syndromes. *J. Biol. Chem.* 2017;292(11):4593–4601.
48. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJE. The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 2015;10(6):845–858.
49. Kircher M, Witten DM, Jain P, et al. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat. Genet.* 2014;46(3):310–315.
50. Karczewski KJ, Francioli LC, Tiao G, et al. The mutational constraint spectrum quantified from variation in 141,456 humans. *bioRxiv*. 2020;531210.
51. Alirezaie N, Kernohan KD, Hartley T, Majewski J, Hocking TD. ClinPred: Prediction Tool to Identify Disease-Relevant Nonsynonymous Single-Nucleotide Variants. *Am. J. Hum. Genet.* 2018;103(4):474–483.
52. Sorror ML, Maris MB, Storb R, et al. Hematopoietic cell transplantation (HCT)-specific comorbidity index: a new tool for risk assessment before allogeneic HCT. *Blood*. 2005;106(8):2912–2919.
53. Tesmer VM, Smith EM, Danciu O, Padmanaban S, Nandakumar J. Combining conservation and species-specific differences to determine how human telomerase binds telomeres. *Proc. Natl. Acad. Sci. U. S. A.* 2019;
54. Banik SSR, Guo C, Smith AC, et al. C-terminal regions of the human telomerase catalytic subunit essential for in vivo enzyme activity. *Mol. Cell. Biol.* 2002;22(17):6234–6246.
55. Gadalla SM, Sales-Bonfim C, Carreras J, et al. Outcomes of allogeneic hematopoietic cell transplantation in patients with dyskeratosis congenita. *Biol. Blood Marrow Transplant.* 2013;19(8):1238–1243.
56. Yabe M, Yabe H, Hattori K, et al. Fatal interstitial pulmonary disease in a patient with dyskeratosis congenita after allogeneic bone marrow transplantation. *Bone Marrow Transplant.* 1997;19(4):389–392.
57. Barbaro P, Vedi A. Survival after Hematopoietic Stem Cell Transplant in Patients with Dyskeratosis Congenita: Systematic Review of the Literature. *Biol. Blood Marrow Transplant.* 2016;22(7):1152–1158.
58. Fioredda F, Iacobelli S, Korthof ET, et al. Outcome of haematopoietic stem cell transplantation in dyskeratosis congenita. *Br. J. Haematol.* 2018;183(1):110–118.
59. Feurstein S, Zhang L, DiNardo CD. Accurate germline RUNX1 variant interpretation and its clinical significance. *Blood Adv.* 2020;4(24):6199–6203.
60. Seimiya H, Sawada H, Muramatsu Y, et al. Involvement of 14-3-3 proteins in nuclear localization of telomerase. *EMBO J.* 2000;19(11):2652–2661.
61. Laprade H, Querido E, Smith MJ, et al. Single-Molecule Imaging of Telomerase RNA Reveals a Recruitment-Retention Model for Telomere Elongation. *Mol. Cell*. 2020;79(1):115–126.e6.
62. Zaug AJ, Podell ER, Nandakumar J, Cech TR. Functional interaction between telomere protein TPP1 and telomerase. *Genes Dev.* 2010;24(6):613–622.
63. Chu TW, D'Souza Y, Autexier C. The Insertion in Fingers Domain in Human Telomerase Can Mediate Enzyme Processivity and Telomerase Recruitment to Telomeres in a TPP1-

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Dependent Manner. *Mol. Cell. Biol.* 2016;36(1):210–222.

64. Luo X, Feurstein S, Mohan S, et al. ClinGen Myeloid Malignancy Variant Curation Expert Panel recommendations for germline RUNX1 variants. *Blood Adv.* 2019;3(20):2962–2979.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Acknowledgements

This work was supported by the National Institutes of Health grants K08CA204734 (R.C.L.), RC2DK122533 (R.C.L.), R01AG050509 (J.N.), and R01GM120094 (J.N.), the Aplastic Anemia & MDS International Foundation (R.C.L.), the Jim and Lois Champy Fund (R.C.L.), the Anna Fuller Fund (R.C.L.), American Cancer Society Research Scholar grant RSF-17-037-01-DMC (J.N.), National Institutes of Health (NIH)/National Heart, Lung, and Blood Institute fellowship training grant 2T32HL116324-06 (C.R.R.), the Sigrid Juselius Foundation (M.M.), the Maud Kuistila Memorial Foundation (M.M.), the Väre Foundation for Pediatric Cancer Research (M.M.) the Orion Research Foundation (M.M.), T32-CA009172 (F.D.T.) the Jock and Bunny Adams Education and Research Fund (J.H.A.), NIH/National Institute of Diabetes and Digestive and Kidney Diseases 1R01DK107716 (S.A.), and the Dana-Farber/Harvard Cancer Center core grant from the National Cancer Institute 5P30 CA006516 (R.R. and D.N.). The Center for International Blood and Marrow Transplant Research is supported primarily by Public Health Service Grant/Cooperative Agreement 5U24CA076518 from the National Cancer Institute, the National Heart, Lung and Blood Institute, and the National Institute of Allergy and Infectious Diseases (W.S., S.R.S. and Z.-H.H.); grant/cooperative agreement 4U10HL069294 from the National Institutes of Health, National Heart, Lung and Blood Institute and National Cancer Institute; Health Resources and Services Administration contract HSH250201200016C; and grants N00014-18-1-2850 and N0014-19-1-2888 from the Office of Naval Research. The views expressed in this article do not reflect the official policy or position of the National Institutes of Health, the Department of the Navy, the Department of Defense, Health Resources and Services Administration or any other agency of the US Government.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Authorship Contributions

M.M., C.R.R., and R.C.L. designed the study, analyzed data, and wrote the manuscript; C.J.G., M.M.C., E.H.O., and I.D.V. performed sequencing; D.N. and R.R. curated clinical data, conceived the statistical plan, and performed statistical analysis; C.R.R., H.Q.R, and J.E.G performed consensus variant classification. V.T., S.P., J.N. performed structural modeling, analyzed data, and contributed to research discussion; F.D.T. and D.K. performed cloning and generated cell lines; W.S., S.R.S., and Z.-H.H. curated clinical data, clinical data collection and quality assurance, and contributed to research discussion; S.A. C.C., and J.H.A. interpreted data and contributed to research discussion; and all authors reviewed the manuscript during its preparation and approved the submission.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Disclosure of Conflicts of Interest

D.N. has received research support from Pharmacyclics and owns stock in Madrigral Pharmaceuticals. M.M. has received honoraria from Celgene and Sanofi. R.C.L. has received research support from Jazz Pharmaceuticals and consulting fees from Takeda Pharmaceuticals and bluebird bio.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1 *TERT* rare variant characteristics

TERT rare variant	Structural domain	gnomAD max popAF	ACMG/AMP Classification	Sherloc Classification	CADD PHRED score	ClinVar Accession Number
p.C76S	TEN	0	VUS	VUS	10.59	
p.V84M	TEN	0	VUS	VUS	22.8	
p.G110A	TEN	0	VUS	VUS	15.96	
p.G135R	TEN	0.0008400	VUS	VUS	16.84	VCV000410665
p.R263H	Linker	0	VUS	VUS	3.266	
p.H296P	Linker	0.0002000	VUS	VUS	4.311	VCV000268080
p.G306S	Linker	0	VUS	VUS	4.83	
p.V435L	TRBD	0	VUS	VUS	6.863	
p.V461E	TRBD	0	VUS	VUS	26.5	
p.E484K	TRBD	0	VUS	VUS	2.57	VCV000955018
p.R485C	TRBD	0.0000638	VUS	VUS	18.03	
p.A532T	TRBD	0.0000265	VUS	VUS	10.15	VCV000581635
p.T567M	TRBD	0	VUS	VUS	17.36	
p.K570R	TRBD	0	VUS	VUS	23.7	
c.1770-2A>G	TRBD	0	VUS	VUS	26.2	
p.R622H	RTD	0	VUS	VUS	24.9	
p.S663G	RTD	0	VUS	VUS	13.49	
p.R669W	RTD	0.0000531	VUS	VUS	22.5	VCV000539196
p.R698Q	RTD	0	VUS	VUS	23.5	
p.G715D	RTD	0	VUS	VUS	24.3	
p.Q722R	RTD	0	VUS	VUS	23.7	VCV000471853
p.V741L	RTD	0.0002000	VUS	VUS	14.03	VCV000652891
p.L766S	RTD	0	VUS	VUS	22.4	
p.P771L	RTD	0	VUS	VUS	24.1	
p.V777M	RTD	0	VUS	VUS	19.9	VCV000436985
p.G847S	RTD	0.00006482	VUS	VUS	24.4	
p.D848N	RTD	0.0000089	VUS	VUS	22.8	
p.R865C	RTD	0.0000240	LP	LP	24.6	VCV000986922
p.V867M	RTD	0.0000000	VUS	VUS	22.6	VCV000242683
p.T917M	RTD	0.0001240	VUS	VUS	21.5	VCV000857994
p.R951W	CTE	0.000008833	VUS	VUS	20.9	VCV000836202
p.S984R	CTE	0.000008828	VUS	VUS	15.71	
p.L994F	CTE	0.00002716	VUS	VUS	22.7	VCV000580043
p.A1014P	CTE	0	VUS	VUS	24.5	
p.S1041F	CTE	0	VUS	VUS	17.63	
p.R1086C	CTE	0	VUS	VUS	15.92	
p.R1086H	CTE	0.000839	VUS	VUS	20.6	VCV000242237
p.V1090M	CTE	0.0003000	VUS	VUS	15.67	VCV000012733
p.R1105W	CTE	0	VUS	VUS	22.2	VCV000939229
p.T1110M	CTE	0.0002000	VUS	VUS	12.05	VCV000039122

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 2. Patient characteristics by *TERT* rare variant status

	No *TERT* rare n = 1473	*TERT* rare n = 41	p-value
Patient-related variables			
Age at transplantation, median (range), years	59 (0 - 77)	52 (14 - 72)	0.03^
Female sex, n (%)	591 (40)	11 (27)	0.11†
Karnofsky performance status score < 90, n (%)	403 (27)	16 (39)	0.15†
Hematopoietic cell transplant comorbidity index			0.15‡
0	255 (25)	3 (11)	
1-2	247 (24)	8 (29)	
3	535 (52)	17 (61)	
Missing	436	13	
Disease-related variables			
Hemoglobin, median (IQR), g/dL	9.4 (8.1-11.2)	9.9 (8.6-11.1)	0.26^
Platelet count, median (IQR), x 10^9/L	72 (30-147)	72 (37-115)	0.87^
Absolute neutrophil count, median (IQR), x 10^9/L	1.1 (0.5-2.3)	1.3 (0.5-2.6)	0.63^
Bone marrow blasts at transplant, median (IQR), %	3 (1-6)	1 (0-5)	0.03^
Prior MDS-directed therapy, n (%)	861 (58)	24 (59)	0.99†
Therapy-related MDS, n (%)	305 (21)	6 (15)	0.43†
Somatic mutations			
ASXL1	289 (20)	8 (20)	0.99†
U2AF1	119 (8)	8 (20)	0.02†
TP53	282 (19)	7 (17)	0.84†
ETV6	57 (4)	5 (12)	0.02†
RUNX1	169 (11)	5 (12)	0.81†
PPM1D	84 (6)	4 (10)	0.29†
Transplant-related variables			
Conditioning regimen, n (%)			0.10†
Myeloablative	765 (52)	24 (59)	
Reduced intensity	565 (39)	17 (41)	
Nonmyeloablative	130 (9)	0 (0)	
Missing	13	0	
Donor type, n (%)			0.85†
Matched, Related	176 (12)	5 (12)	
Matched, Unrelated	837 (57)	26 (63)	
Mismatched	289 (20)	7 (17)	
Cord Blood	171 (12)	3 (7)	
Graft type, n (%)			0.88†
Bone marrow	215 (15)	6 (15)	
Peripheral blood stem cells	1,082 (73)	32 (78)	
Cord Blood	165 (11)	3 (7)	
Other	11 (1)	0 (0)	

^Wilcoxon rank-sum test,
†Fisher's exact test,
‡Cochran-Armitage trend test
Peripheral blood counts and bone marrow blast counts at time of transplantation
IQR, interquartile range

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 1.

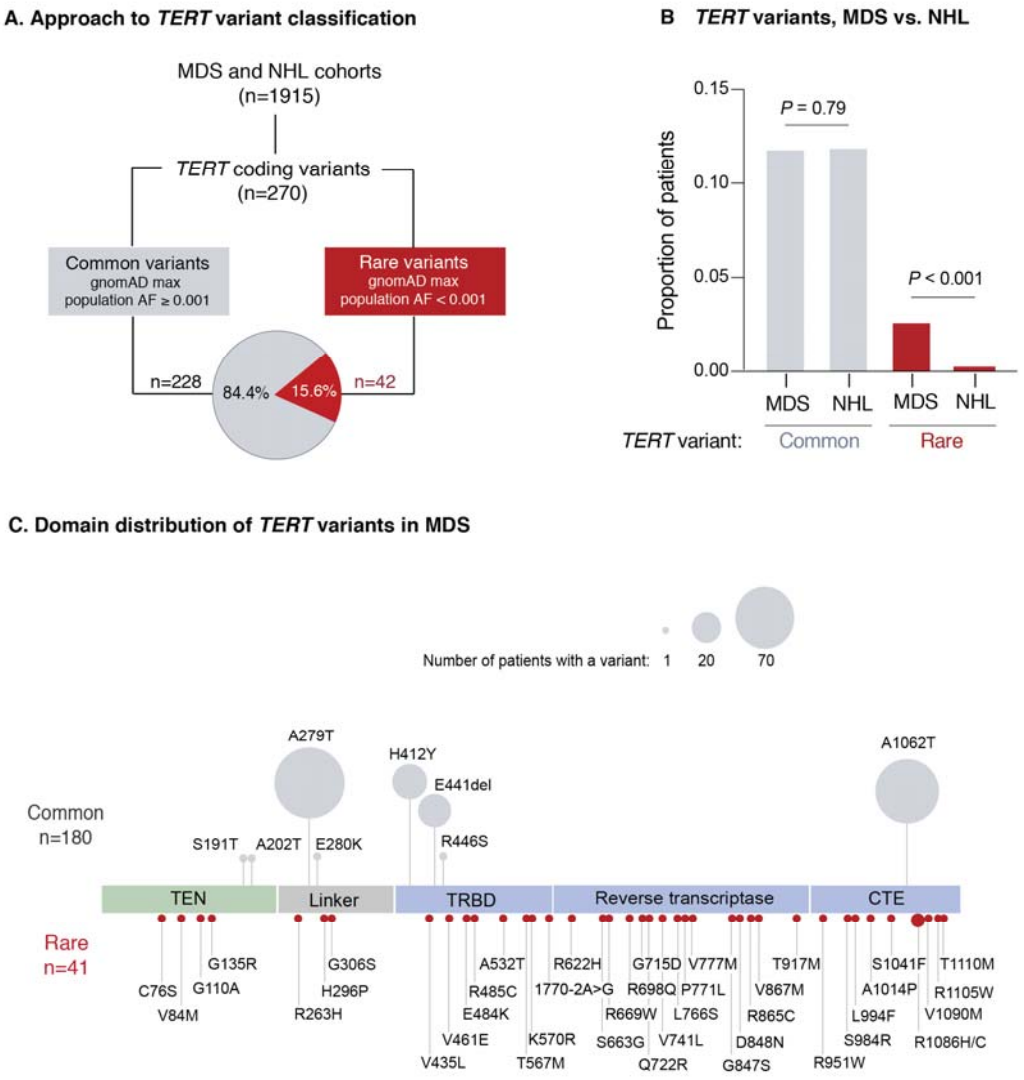


Figure 1. *TERT* variants in MDS and NHL. A) Classification approach of nonsynonymous *TERT* variants identified in the MDS and NHL cohorts. B) Frequency of *TERT* common and *TERT* rare variants within the MDS and NHL cohorts. C) Domain distribution of *TERT* variants within the MDS cohort. *TERT* common variants (n= 180 variants) and rare variants (n=40 variants among 41 patients) are located above and below the coding region, respectively. The size of each ball is proportional to the number of patients with that variant. *TERT* rare variants are colored in red and *TERT* common variants in gray.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 2.

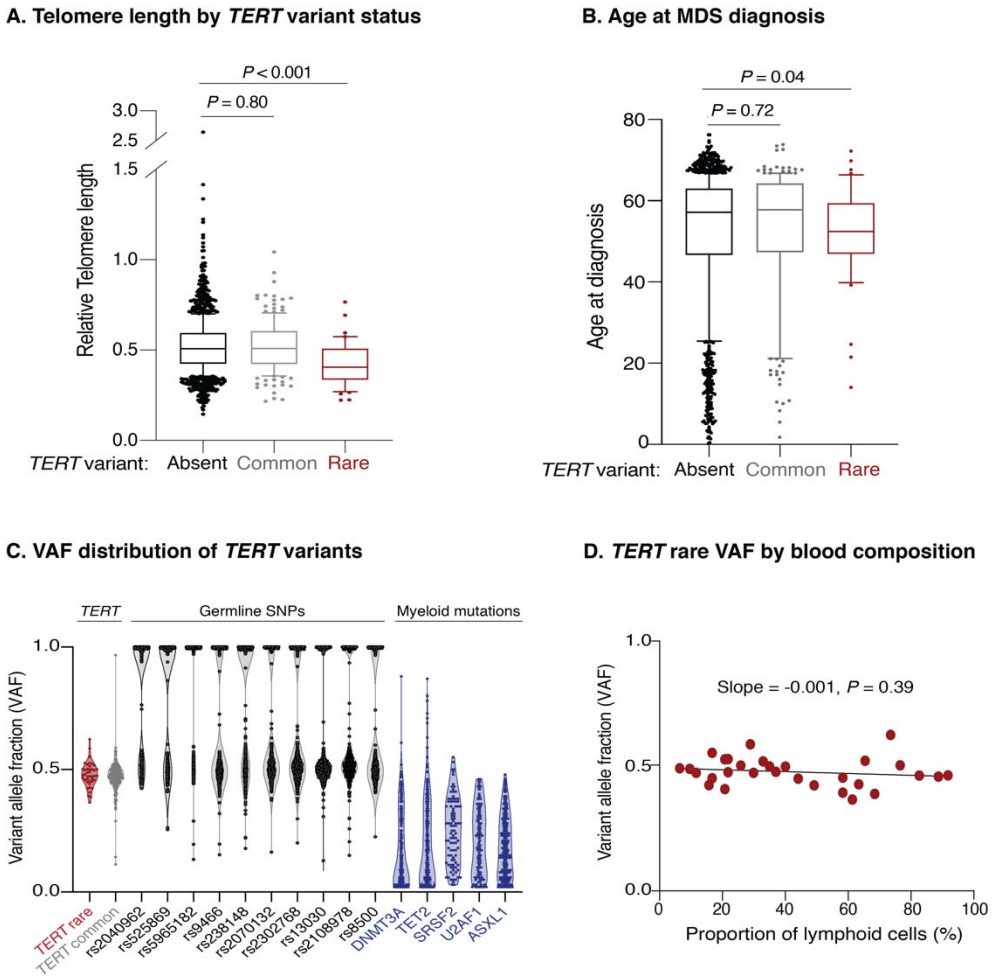
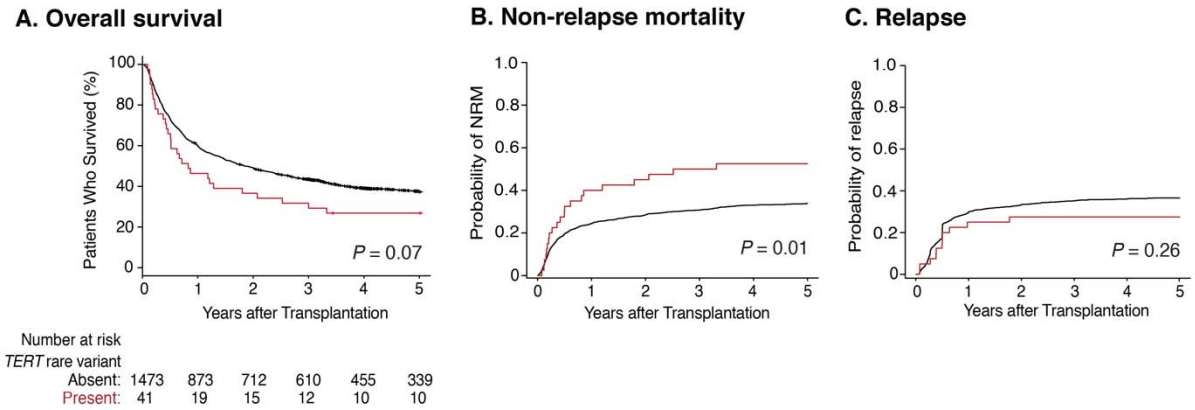


Figure 2. Association of TERT variants with telomere length and age at MDS diagnosis.

A) Pre-transplant whole blood relative telomere length by *TERT* variant status. B) Age at MDS diagnosis by *TERT* variant status. *TERT* variant groups are labelled as follows: no *TERT* variant (black), *TERT* common variant (gray), and *TERT* rare variant (red). C) Variant allele fraction distribution of *TERT* variants, germline SNPs (black), and myeloid mutations (blue). D) Variant allele fraction distribution of *TERT* rare variants as a function of the proportion of blood lymphocytes. Linear regression slope with p-value is shown.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 3.



D. Multivariable models of transplant outcomes

		Cox regression: OS		Competing risks regression: NRM		Competing risks regression: relapse	
	No. of patients	HR (95% CI)	P	HR (95% CI)	P	HR (95% CI)	P
***TERT* rare variant status**							
Absent (reference)	1473						
Present	41	1.50 (1.04, 2.17)	0.03	1.75 (1.13, 2.71)	0.01	0.78 (0.42, 1.46)	0.44
Recipient age							
10 year increase	1514	1.23 (1.16, 1.30)	<0.001	1.23 (1.14, 1.33)	<0.001	0.98 (0.91, 1.04)	0.48
Karnofsky Performance Score							
90-100 (reference)	817						
10-80	419	1.27 (1.10, 1.48)	0.002	1.23 (1.02, 1.53)	0.03	0.99 (0.80, 1.22)	0.94
Missing	278	1.05 (0.87, 1.27)	0.59	0.95 (0.73, 1.24)	0.72	1.14 (0.90, 1.45)	0.27
HCT-CI score group							
0 (reference)	258						
1-2	255	1.29 (1.01, 1.65)	0.04	1.15 (0.83, 1.58)	0.40	1.20 (0.88, 1.63)	0.25
3 or above	552	1.47 (1.19, 1.83)	<0.001	1.32 (0.99, 1.76)	0.06	1.07 (0.81, 1.40)	0.64
Missing	449	1.35 (0.86, 2.11)	0.19	0.78 (0.41, 1.47)	0.44	1.86 (1.12, 3.07)	0.02

Figure 3. Transplant outcomes by *TERT* rare variant status. A) Kaplan-Meier curve for overall survival. B) Cumulative incidence curves for non-relapse mortality. C) Cumulative incidence curves for relapse. *TERT* rare variants are colored in red and patients without a *TERT* rare variant are colored in black. D. Multivariable models of overall survival (left), NRM (middle), and relapse (right).

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 4.

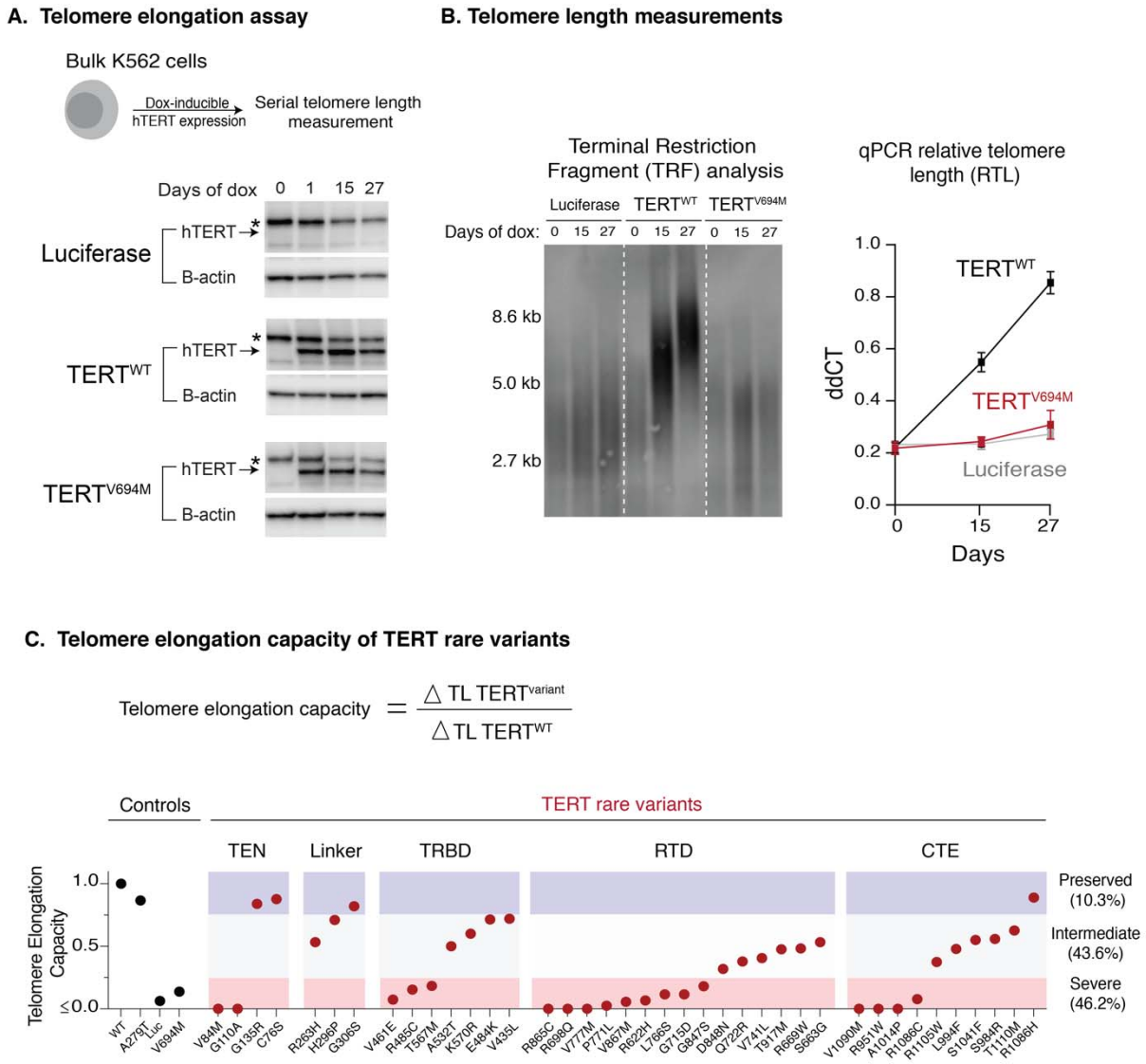


Figure 4. Functional characterization of *TERT* rare variants. A) Cell-based telomere elongation assay in isogenic bulk K562 cell lines with doxycycline-inducible TERT expression. hTERT expression throughout the experiment is shown for control conditions: luciferase, $TERT^{WT}$, $TERT^{V694M}$. hTERT band (~127kDa) is labelled with an arrow and the asterisk corresponds to a non-specific band seen in all conditions. B) Telomere length measurements by terminal restriction fragment analysis and qPCR for luciferase (gray), $TERT^{WT}$ (black) $TERT^{V694M}$

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

(red). C) Telomere elongation capacity of *TERT* rare variants normalized to wild-type TERT rare shown in ranked order grouped by structural domain. Control conditions are colored in black and *TERT* rare variants in red.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Figure 5.

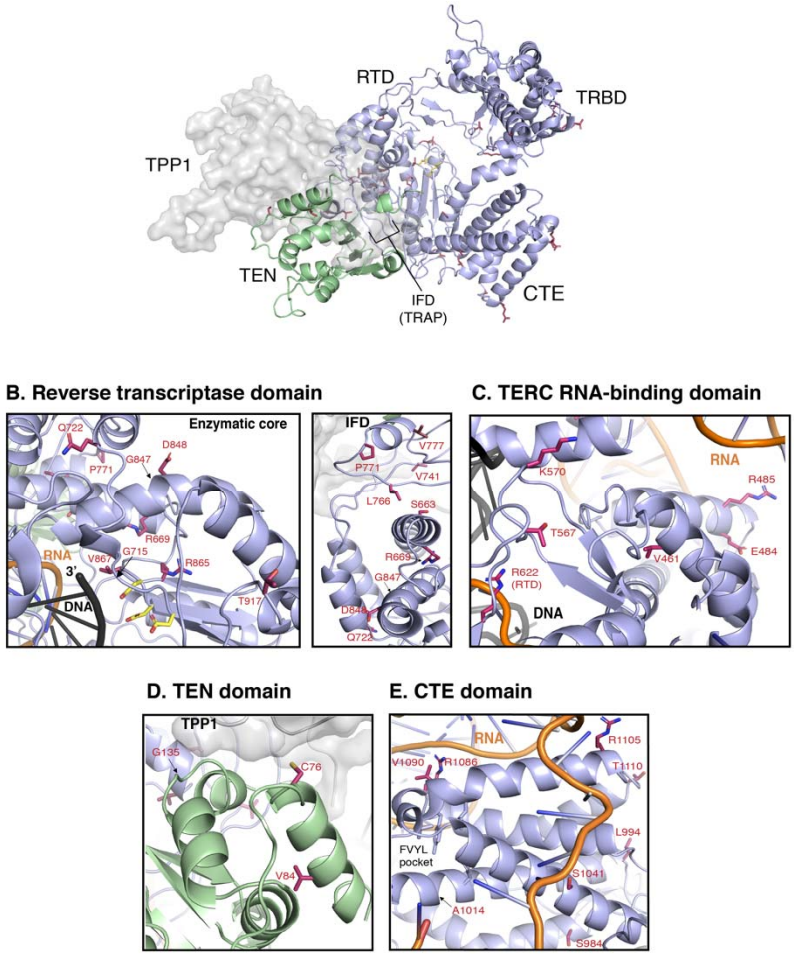


Figure 5. Structural analysis of *TERT* rare variants. A) Human TERT homology model. The ring, formed by the TRBD, RTD, and CTE domain is colored in purple and the TEN domain in green. Panels B, C, D, and E show the *TERT* rare variants within the RTD (enzymatic core and IFD regions), TRBD, TEN, and the CTE domains (including the FVYL pocket), respectively. Note that R622 belongs to the RTD but is displayed in panel C due to its proximity to the TRBD. The side-chains for the residues mutated in the MDS rare variants (carbon atoms colored crimson) and the catalytic residues in the active site (carbon atoms colored yellow) are shown in stick representation. The modeled regions of TERC are in orange while the DNA substrate is depicted in black.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Table of Contents

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Methods

DNA sequencing

Amplicon libraries were prepared with an Ion AmpliSeq Custom panel by using the Ampliseq Library Kit Plus (Thermo Fisher Scientific, USA) according to the manufacturer's protocol. These amplicons were partially digested before ligation of adapters and multiplex barcodes. Ligation products were purified with AMPure XP beads (Beckman Coulter, USA) and quantified using the Ion Library TaqMan Quantitation Kit (Thermo Fisher Scientific, USA). Libraries were normalized to 30 pM concentration and multiplexed in batches of no more than 96 samples for further processing. Templating and Ion 530 chip loading was performed on the Ion Chef (Thermo Fisher Scientific, USA) followed by sequencing on the Ion S5 (Thermo Fisher Scientific, USA) according to the manufacturer's instructions. Samples were sequenced to at least 150x average read depth. Raw reads were aligned with the TMAP alignment package and variants called using the VariantCaller plugin, both from the Torrent Suite software (Thermo Fisher Scientific, USA). Variant annotation was done using Annovar[47]. Intronic and synonymous variants and variants with less than 5 alternate reads or less than 10 total reads were excluded from the analysis.

TERT variant classification

TERT variants were classified as "common" or "rare" using a maximum population allele frequency in gnomAD (v2.1.1) of 0.001 in any reference population. *TERT* common and rare variants were classified by consensus as benign (B), likely benign (LB), variant of unknown significance (VUS), likely pathogenic (LP) or pathogenic (P) according to ACMG/AMP[33] and Sherloc[34] guidelines through manual curation. *In silico* prediction models Combined Annotation-Dependent Depletion (CADD)[50] and ClinPred[51] were used to estimate likelihood of pathogenicity of each variant. A CADD PHRED score of 20 or greater indicates a SNV is in the top 1% of deleterious variants across all possible SNVs in human genome reference (

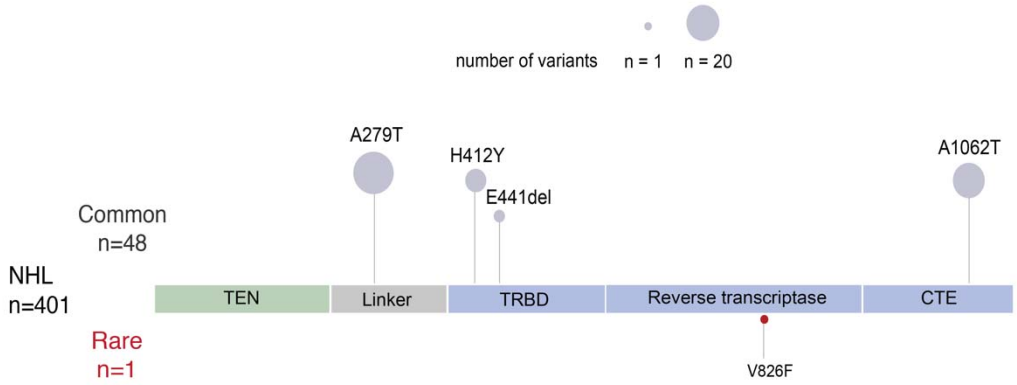
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

cadd.gs.washington.edu). A ClinPred score >0.5 is considered a prediction of pathogenicity for a given variant (sites.google.com/site/clinpred/).

hTERT structural homology model

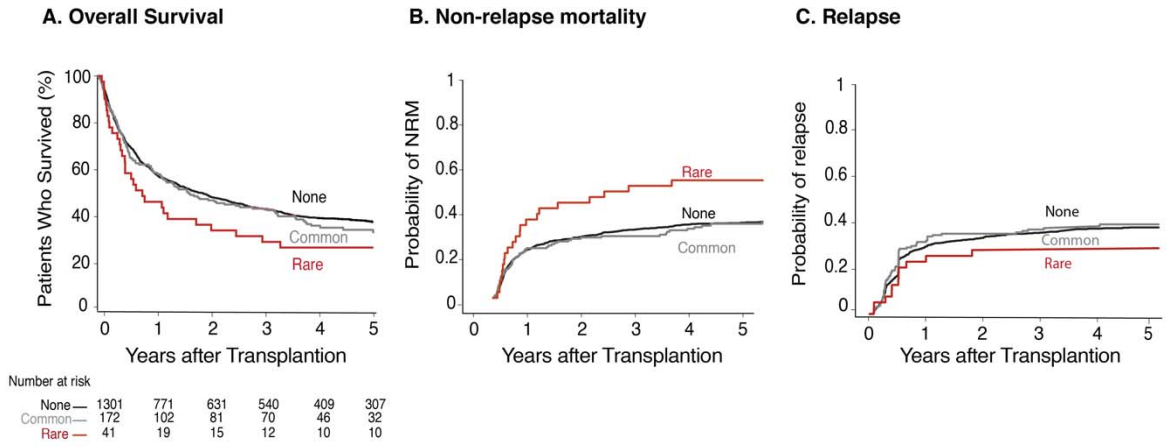
The full human TERT TEN domain, RTD, and the template-primer complex were modeled using the *Tetrahymena thermophila* telomerase holoenzyme cryo-EM structure (PDB accession code: 6D6V). The human TERT RBD domain bound to TR was modeled using the crystal structure of *Takifugu rubripes* TERT RBD-TR CR4/5 complex (PDB accession code: 4LMO) as a template. The solved crystal structures of human TERT CTE (PDB accession code: 5UGW) and human TPP1-OB (PDB: 2I46) were used as such. The full model of human telomerase-TPP1-OB was assembled from these individual homology models by using the *Tetrahymena thermophila* telomerase holoenzyme cryo-EM reconstruction (PDB accession code: 6D6V) as a scaffold. The human TERT TEN, RT (and associated template-primer duplex), RBD (and associated CR4/5 RNA), and CTE were superimposed on their counterparts in the *Tetrahymena thermophila* telomerase holoenzyme reconstruction in Pymol using the “Align” command to assemble the human telomerase part of the homology model. Human TPP1-OB was superimposed on the p50 subunit of the *Tetrahymena thermophila* telomerase holoenzyme reconstruction to complete the human telomerase-TPP1 homology model. While TERT and TPP1 amino acids in the homology model correspond to sequences found in the human polypeptides, the original sequences and structures were retained for TR (*Tetrahymena thermophila* TR sequence for the template region and *Takifugu rubripes* TR sequence for the CR4/5 domain) and primer (*Tetrahymena thermophila* telomeric DNA primer sequence) regions in the homology model. No energy minimization or other refinement of the homology model was performed after assembly from individual components.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



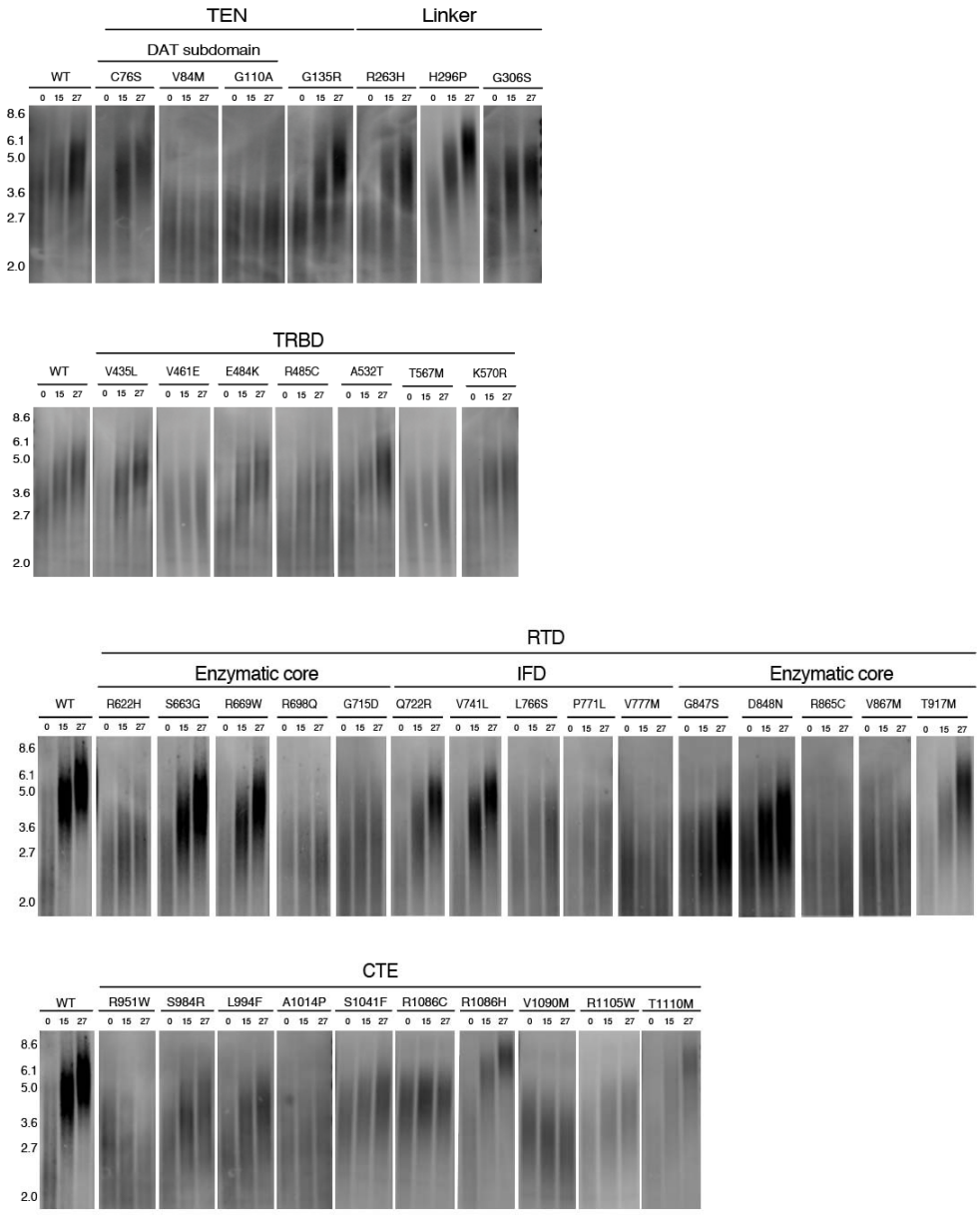
Supplemental Figure 1. *TERT* variants in NHL. Domain distribution of *TERT* variants within the NHL cohort. *TERT* common variants (n = 48) and rare variant (n =1) are located above and below the coding region, respectively. The size of each ball is proportional to the number of patients with that variant. The single *TERT* rare variant is colored in red and *TERT* common variants in gray.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplemental Figure 2. Transplant outcomes by *TERT* variant status. A) Kaplan-Meier curve of overall survival. B) Cumulative incidence curve for non-relapse mortality. C) Cumulative incidence curve for relapse. Patients with a *TERT* rare variants, *TERT* common variants, or no *TERT* variant are colored in red, gray, and black, respectively.

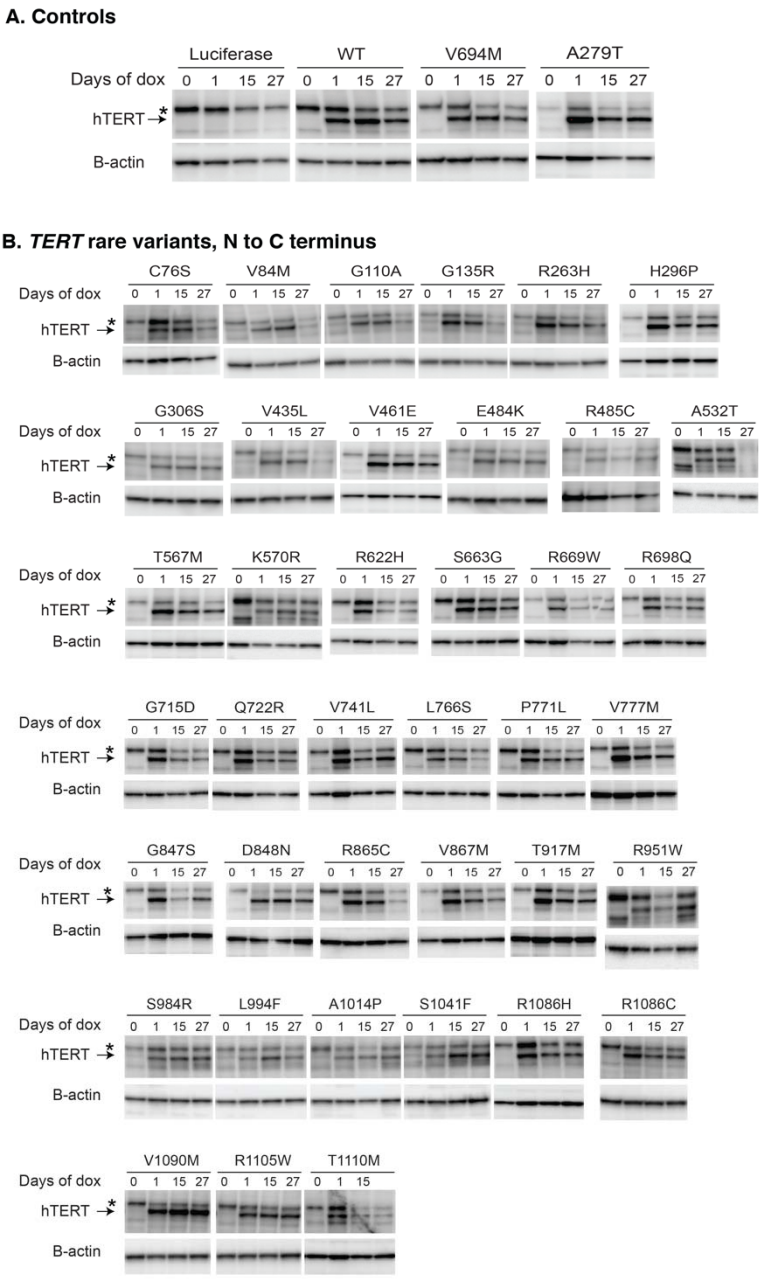
bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplemental Figure 3 Terminal restriction fragment analysis of *TERT* rare variants.

TRF images of 2ug digested genomic DNA extracted at Day 0, 15, and 27 for each *TERT* variant run on 1% agarose gel. Variants are arranged by structural domain and depict data from multiple gels.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

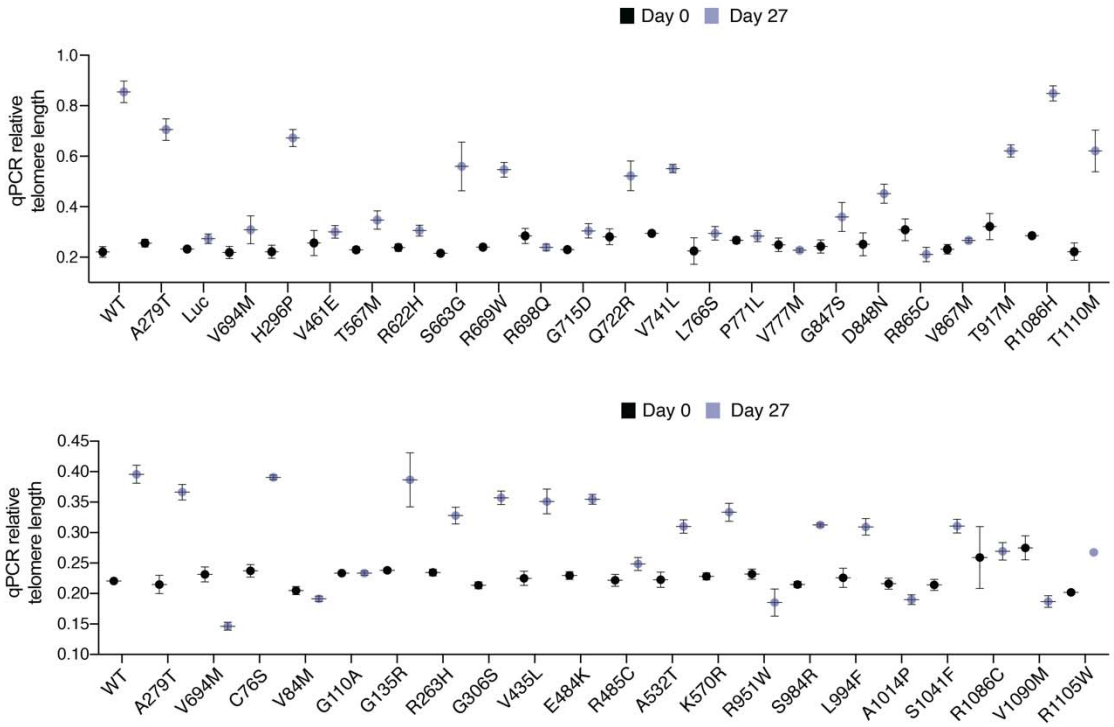


Supplemental Figure 4. TERT western blot summary. Induced expression of hTERT was measured at Day 0 prior to starting doxycycline and Day 1, 15, and 27 during the experiment. Arrow corresponds to the expected molecular weight of hTERT (~127 kDa) and asterisks corresponds to non-specific band seen in all conditions. A) Control conditions: luciferase,

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$TERT^{WT}$, $TERT^{V694M}$, and the common variant $TERT^{A279T}$. B) *TERT* rare variants from N to C terminus.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.



Supplemental Figure 5. qPCR relative telomere length summary of *TERT* rare variants.

Graphs of relative telomere length measurements for each *TERT* rare variant at the beginning (Day 0 - black) and end (Day 27 - blue) of the experiment. Error bars correspond to standard deviation of triplicate values. qPCR measurements for all variants were performed in two batches.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Table 1

TERT variant	Forward Primer	Reverse Primer
C76S	CAGGTGTCCTCCCTGAAGGAGCTGGTGG	GCGGAAGGAGGGGGCGGC
V84M	GGTGGCCCGAATGCTGCAGAG	AGCTCCTTCAGGCAGGACACC
G110A	GCCCGCGGGGCCCCCCCCGAG	CCCGTCCAGCAGCGCGAAGCCG
G135R	GCGGGGGAGCAGGGCGTGGGG	AGTGCGTCGGTCACCGTGTTGGG
R263H	GGCAGGACGCATGGACCGAGTG	CGGGTGGGCCCAGGACCC
A279T	TGCCAGACCCACCGAAGAAGC	GGTGACACCACACAGAAAC
H296P	CGCCACTCCCCCCCATCCGTG	CGTGCCAGAGAGCGCACC
G306S	GCACCACGCGAGCCCCCCATC	TGGCGGCCCACGGATGGG
V435L	CCAGGGCTCTCTGGCGGCCCC	GGCTTCTCCCGGGCACAG
V461E	CCCTGGCAGGAGTACGGCTTC	GCTGCTGTGCTGGCGGAG
E484K	CAGGCACAACAAACGCCGCTTCC	GAGCCCCAGAGGCCTGGG
R485C	GCACAACGAATGCCGCTTCCTC	CTGGAGCCCCAGAGGCCT
A532T	TGTTCCGGCCACAGAGCACCG	CAGCCAACCCCTGGGCTC
T567M	ACGGAGACCATGTTTCAAAAG	GACATAAAAGAAAGACCTGAG
K570R	ACGTTTCAAAGGAACAGGCTC	GGTCTCCGTGACATAAAAG
R622H	TCCAGACTCCACTTCATCCCCAAGCCTGAC	CGTCAGCAGGGCGGGCCT
S663G	GGCACTGTTCGGCGTGCTCAA	TTCACCCTCGAGGTGAGAC
R669W	CAACTACGAGTGGGCGCGGCG	AGCACGCTGAACAGTGCCTTC
V694M	GCGCACCTTCATGCTGCGTGT	CAGGCCCTGTGGATATCGTC
R698Q	CTGCGTGTGCAGGCCCAGGAC	CACGAAGGTGCGCCAGGC
G715D	GATGTGACGGACGCGTACGAC	CACCTTGACAAAGTACAGCTC
Q722R	ACCATCCCCCGGGACAGGCTC	GTCGTACGCGCCCGTCAC
V741L	CACGTACTGCCTGCGTCGGTA	TTCTGGGGTTTGATGATGC
L766S	GTCTCTACCTCGACAGACCTCC	GTGGCTCTTGAAGGCCTT
P771L	GACCTCCAGCTGTACATGCGA	TGTCAAGGTAGAGACGTGG
V777M	GCGACAGTTCATGGCTCACCT	ATGTACGGCTGGAGGTCT
G847S	CCTGTGCTACAGCGACATGGAG	CTGCAGAGCAGCGTGGAG
D848N	GTGCTACGGCAACATGGAGAACAAG	AGGCTGCAGAGCAGCGTG
R865C	GCTGCTCCTGTGTTTGGTGGATGATTTC	CCGTCCCGCCGAATCCCC
V867M	CCTGCGTTTGATGGATGATTTCTTGTTGGTGACACC	AGCAGCCCGTCCCGCCGA
T917M	CTGGGTGGCATGGCTTTTGTT	GGCCTCGTCTTCTACAGG
R951W	CAGCTATGCCTGGACCTCCAT	GAGTAGTCGCTCTGCACC
S984R	AGTGTCACAGGCTGTTTCTGGATTTG	TCAGCCGCAAGACCCCAA
L994F	GGTGAACAGCTTCCAGACGGT	TGCAAATCCAGAAACAGGC
A1014P	CAGGTTTCACCCATGTGTGCTGCAG 3	TACGCCTGCAGCAGGAGG
S1041F	GACACGGCCTTCCTCTGCTAC	AGAGATGACGCGCAGGAAAAATG
R1086C	GACTCGACACTGTGTCACCTA	AGCTTGAGCAGGAATGCT
R1086H	ACTCGACACCATGTCACCTAC	CAGCTTGAGCAGGAATGC
V1090M	TGTCACCTACATGCCACTCCT	CGGTGTCGAGTCAGCTTG
R1105W	GCAGCTGAGTTGGAAGCTCCC	GTCTGGGCTGTCCTGAGTG
T1110M	CTCCCGGGGATGACGCTGACT	CTTCCGACTCAGCTGCGTC

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Table 2

Ch 5 genomic position	Ref	Alt	*TERT* rare variant	ACMG/AMP Criteria	Sherloc Criteria	ClinPred score	PMID Reference
1294774	C	G	p.C76S	PM2; BP4	EV01035	0.05921087	
1294751	C	T	p.V84M	PM2; PP3	EV01035	0.90967768	
1294672	C	G	p.G110A	PM2; PP3, PP5	EV01035, EV0061, EV0053	0.42297935	26024875
1294598	C	T	p.G135R	PM2	EV0161	0.02669442	
1294213	C	T	p.R263H	PM2; BP4	EV01035, EV0126	0.20593388	
1294114	T	G	p.H296P	PM2; BP4	EV0161, EV0126	0.05764545	
1294085	C	T	p.G306S	PM2; BP4	EV01035, EV0126	0.04076032	
1293698	C	G	p.V435L	PM1, PM2; BP4	EV01035, EV0172 EV0126	0.12568937	
1293619	A	T	p.V461E	PM1, PM2; PP3	EV01035, EV0172 EV0122	0.99402605	
1293551	C	T	p.E484K	PM1, PM2; BP4	EV01035, EV0172 EV0126	0.02799475	
1293548	G	A	p.R485C	PM1, PM2	EV0101, EV0172, EV0024	0.0912887	17460043
1282719	C	T	p.A532T	PM1, PM2; BP4	EV0101, EV0172, EV0126	0.05038688	
1282613	G	A	p.T567M	PM1, PM2; PP3	EV01035, EV0172, EV0053	0.71335232	23538340
1282604	T	C	p.K570R	PM1, PM2, PM5; PP3	EV01035, EV0172, EV0044	0.86071074	
1280455	T	C	c.1770-2A>G	PM2	EV01035, EV0172	n/a	
1280358	C	T	p.R622H	PM1, PM2; PP3	EV01035, EV0172, EV0122	0.9958812	25346280
1279549	T	C	p.S663G	PM1, PM2	EV01035, EV0172	0.14513995	
1279531	G	A	p.R669W	PM1, PM2; PP3	EV0101, EV0172, EV0122	0.94496411	
1279443	C	T	p.R698Q	PM1, PM2; PP3	EV01035, EV0172, EV0081	0.45770326	22664374
1278898	C	T	p.G715D	PM1, PM2; PP3	EV01035, EV0172	0.99102062	
1278877	T	C	p.Q722R	PM1, PM2; PP3	EV01035, EV0172	0.97397363	
1278821	C	A	p.V741L	PM1, PM2	EV0101, EV0172	0.07110612	
1272385	A	G	p.L766S	PM1, PM2; PP3	EV01035, EV0172	0.5904355	
1272370	G	A	p.P771L	PM1, PM2; PP3	EV01035, EV0172	0.99044496	
1272353	C	T	p.V777M	PM1, PM2, PM5, PP3	EV01035, EV0172	0.92925555	25741868
1268678	C	T	p.G847S	PM1, PM2; PP3	EV0101, EV0172	0.99762005	
1268675	C	T	p.D848N	PM1, PM2; PP3	EV0101, EV0172	0.9248184	
1266640	G	A	p.R865C	PM1, PM2; PP3, PP5	EV0101, EV0172, EV0024, EV0122	0.99269283	17460043
1266634	C	T	p.V867M	PM1, PM2; PP3	EV01035, EV0172, EV0024	0.82887769	20502709
1264612	G	A	p.T917M	PM1, PM2; PP3	EV0101, EV0172	0.27089664	
1260708	G	A	p.R951W	PM2; PP1, PP3	EV0101, EV0051, EV0024, EV0122	0.67566538	20502709; 27540018
1260607	G	C	p.S984R	PM2; BP4	EV0101	0.31716758	

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1258765	G	A	p.L994F	PM2	EV0101	0.44978669	
1255519	C	G	p.A1014P	PM2; PP3	EV01035	0.98364198	
1255437	G	A	p.S1041F	PM2; PP3	EV01035	0.87835681	
1254522	G	A	p.R1086C	PM2	EV01035, EV0122	0.96164751	28099038
1254521	C	T	p.R1086H	PM2	EV0161	0.05256126	
1254510	C	T	p.V1090M	PM2; PP5	EV0161, EV0080, EV0024	0.03671261	15814878; 23901009
1253929	G	A	p.R1105W	PM2; PP3	EV01035, EV0221	0.73196268	
1253913	G	A	p.T1110M	PM2; PP6	EV0161, EV0221	0.04365337	17392301

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Table 3

Ch 5 genomic position	Reference	Alternate	*TERT* common variant	Structural domain	gnomAD max popAF	ACMG/AMP criteria	Sherloc criteria	ClinVar Accession Number
1294429	C	G	p.S191T	TEN	0.003646	B	B	VCV000350804
1294397	C	T	p.A202T	TEN	0.00182	B	B	VCV000012729
1294166	C	T	p.A279T	TEN	0.1209	B	B	VCV000039125
1294163	C	T	p.E280K	TEN	0.00277	LB	B	VCV000471904
1293767	G	A	p.H412Y	TRBD	0.0188	B	B	VCV000012730
1293677	0	-	p.E441del	TRBD	0.003621	B	B	VCV000212398
1293665	G	T	p.R446S	TRBD	0.002254	LB	LB	VCV000242216
1254594	C	T	p.A1062T	CTE	0.02149	B	B	VCV000039121

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Table 4

Patient ID	RTL	Age at transplant	*TERT* rare variant	Variant alle fraction (VAF)
131-871-9	0.530985789	59.7	p.C76S	0.5517
632-160-1	0.338146726	52.6	p.V84M	0.500
167-534-0	0.38999983	68.0	p.G110A	0.3882
012-006-6	0.766510528	58.5	p.G135R	0.624
004-659-9	0.574438667	56.9	p.R263H	0.513
695-506-9	0.292339718	25.1	p.H296P	0.4486
193-453-1	0.475952599	72.9	p.G306S	0.4981
108-599-5	0.559849249	59.1	p.V435L	0.4764
006-488-1	0.257154532	49.4	p.V461E	0.526
684-474-3	0.412026455	63.3	p.E484K	0.4905
182-114-2	0.350501483	66.2	p.R485C	0.4228
202-087-6	0.557648152	45.9	p.A532T	0.4974
188-048-6	0.223210607	46.6	p.T567M	0.5269
112-386-1	0.516129796	60.6	p.K570R	0.4749
139-119-5	0.225024512	73.0	c.1770-2A>G	0.4609
199-354-5	0.405217397	52.9	p.R622H	0.422
672-577-7	0.401343382	52.8	p.S663G	0.5024
680-768-2	0.440760004	49.7	p.R669W	0.5181
622-265-0	0.394371706	44.2	p.R698Q	0.5228
675-834-9	0.329298568	45.4	p.G715D	0.3652
138-295-4	0.57269467	55.7	p.Q722R	0.4072
646-745-3	0.411590539	70.6	p.V741L	0.4719
631-993-6	0.32504369	48.2	p.L766S	0.4414
116-369-3	0.335353357	47.1	p.P771L	0.5012
658-186-5	0.343358458	57.0	p.V777M	0.5204
126-140-6	0.291944859	52.3	p.G847S	0.4721
169-536-3	0.336384474	50.9	p.D848N	0.4718
137-528-9	0.317172711	48.4	p.R865C	0.4533
618-645-9	0.462689547	60.3	p.V867M	0.4801
653-872-5	0.418968207	62.0	p.T917M	0.5008
648-041-5	0.356232164	53.0	p.R951W	0.465
694-861-9	0.692395209	49.3	p.S984R	0.4878
637-301-6	0.545545414	39.7	p.L994F	0.5012
618-695-4	0.400868653	47.6	p.A1014P	0.4567
001-142-9	0.595007843	48.4	p.S1041F	0.4733
126-937-5	0.42325787	53.4	p.R1086C	0.4618
216-744-6	0.450950753	67.1	p.R1086H	0.3925

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

223-987-2	0.366835906	14.1	p.R1086H	0.4264
141-314-8	0.501318668	66.9	p.V1090M	0.4502
155-275-4	0.500527712	59.1	p.R1105W	0.5865
671-880-6	0.265103336	22.2	p.T1110M	0.4573

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Table 5

TERT rare variant	Telomere elongation capacity		hTERT homology model structural comments
p.C76S	preserved	0.88	Surface residue within DAT subdomain implicated in TPP1 binding; Ser results in mild change but decreased hydrophobicity; predicted likely tolerated.
p.V84M	severe	<0.0	Core residue in DAT subdomain implicated in TPP1 binding; Met substitution increases bulk but preserves hydrophobicity; predicted likely tolerated
p.G110A	severe	0.00	Not modeled but within DAT subdomain implicated in TPP1 binding; predicted pobably tolerated.
p.G135R	preserved	0.84	Located distal to DAT subdomain; Arg predicted to disrupt helix orientation and impair TPP1 binding; predicted likely pathogenic.
p.R263H	intermediate	0.53	Not modeled
p.H296P	intermediate	0.71	Not modeled
p.G306S	preserved	0.82	Not modeled
p.V435L	intermediate	0.72	Not modeled but likely in unstructured loop at a distance from RNA; predicted tolerated.
p.V461E	severe	0.07	In hydrophobic core; Glu likely disrupts folding; predicted pathogenic.
p.E484K	intermediate	0.71	Acidic residue in basic patch close to RNA; Lys could destabilize the basic helix; Predicted likely pathogenic
p.R485C	severe	0.15	On basic surface close to RNA binding groove; Cys results in loss of positive charge; predicted pathogenic
p.A532T	intermediate	0.50	On a helix that binds CR4/CR5 domain; Thr reduces hydrophobicity; predicted likely tolerated.
p.T567M	severe	0.18	On a hydrophilic loop close to the RNA-DNA duplex; Met increases bulk; predicted likely pathogenic
p.K570R	intermediate	0.60	On a hydrophilic loop close to RNA-DNA duplex; Arg maintains positive charge; predicted likely tolerated.
c.1770-2A>G	n/a	n/a	Not modeled
p.R622H	severe	0.07	Close to the template phosphodiester backbone; His likely disrupts this interaction; predicted pathogenic
p.S663G	intermediate	0.53	Surface residue facing IFD; Gly unlikely to produce significant change in structure; predicted likely tolerated
p.R669W	intermediate	0.48	Residue inserts between two helices; Trp could be accommodated; predicted likely tolerated
p.R698Q	severe	<0.0	Surface residue that forms H-bond to a beta-turn. Gln would allow H-bond; predicted likely tolerated.
p.G715D	severe	0.12	Active site residue; Asp would repel DNA-RNA duplex; predicted pathogenic
p.Q722R	intermediate	0.38	On a IFD bracing helix; Arg adds positive charge; predicted likely pathogenic.
p.V741L	intermediate	0.41	Hydrophobic IFD reside; Leu is minimal change; predicted tolerated.
p.L766S	severe	0.12	Hydrophobic IFD residue facing enzymatic core. Ser reduces hydrophobicity; predicted likely tolerated.
p.P771L	severe	0.03	Surface residue at TEN-IFD junction that could be involved in TPP1 interactions; retained hydrophobicity; predicted likely pathogenic.
p.V777M	severe	<0.0	Residue of IFD helix contacting TEN and TPP1; Met increases bulk; predicted pathogenic.
p.G847S	severe	0.18	Residue forms a kink in a long helix; Ser could disrupt conformation; predicted pathogenic.
p.D848N	intermediate	0.32	Residue faces IFD bracing helices; Asn may disrupt salt bridge with R724; predicted likely pathogenic.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

p.R865C	severe	<0.0	Active side residue; Cys leads to loss of salt bridge with E850; predicted pathogenic.
p.V867M	severe	0.06	Residue in close proximity to DNA; Met increases bulk; predicted likely pathogenic
p.T917M	intermediate	0.48	Surface residue away from any interface; Met increases hydrophobicity; predicted likely tolerated.
p.R951W	severe	<0.0	Not modeled or conserved; predicted likely tolerated.
p.S984R	intermediate	0.56	Surface residue away from any interface; predicted tolerated
p.L994F	intermediate	0.48	Hydrophobic region and Phe will increase bulk; predicted likely tolerated.
p.A1014P	severe	<0.0	Strongly conserved and likely helix-breaking; predicted pathogenic.
p.S1041F	intermediate	0.55	Hydrophilic residue within 14-3-3 binding site on helix involved in RNA binding; predicted likely pathogenic
p.R1086C	severe	0.08	Close to FVYL pocket involved in P6.1 binding; Cys less hydrophilic; predicted likely tolerated.
p.R1086H	preserved	0.89	Close to FVYL pocket involved in P6.1 binding; His predicted tolerated.
p.V1090M	severe	<0.0	Surface residue; Met increases bulk; predicted tolerated but previously determined to be dysfunctional and disease-associated.
p.R1105W	intermediate	0.37	Very close to RNA backbone; Trp results in loss of positive charge; predicted likely pathogenic.
p.T1110M	intermediate	0.63	Poorly conserved surface residue at a distance from RNA; predicted tolerated.

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Table 6

	TERT variant		
	None	Common	Rare
	n = 1301 (86)	n = 172 (11)	n = 41 (3)
Patient-related variable			
Age at transplantation, median (range), years	59 (0 - 77)	59 (5 - 75)	52 (14 - 72)
Female sex, n (%)	519 (40)	72 (42)	11 (27)
Karnofsky performance status score < 90, n (%)	349 (27)	54 (31)	16 (39)
Hematopoietic cell transplan comorbidity index (HCT-CI)			
0	225 (24)	30 (25)	3 (11)
1-2	225 (24)	22 (19)	8 (29)
3	469 (51)	66 (56)	17 (61)
Missing	382	54	13
Disease-related variable			
Hemoglobin, median (interquartile range), g/dL	9.4 (8.1-11.2)	9.2 (8.0-11.0)	9.9 (8.6-11.1)
Platelet count, median (interquartile range), x 10^9/L	73 (30-148)	68 (23-140)	72 (37-115)
Absolute neutrophil count, median (interquartile range), x 10^9/L	1.1 (0.5-2.3)	1.4 (0.5-2.3)	1.3 (0.5-2.6)
Bone marrow blasts at transplant, median (interquartile range), %	3 (1-6)	2 (1-7)	1 (0-5)
IPSS-R cytogenetic risk group, n (%)			
Very good	8 (1)	-	1 (3)
Good	461 (45)	66 (48)	15 (43)
Intermediate	216 (21)	28 (20)	9 (26)
Poor	226 (22)	34 (25)	5 (14)
Very poor	107 (11)	10 (7)	5 (14)
Unknown	283	34	6
IPSS-R group, n (%)			
Very low	101 (10)	14 (11)	4 (13)
Low	249 (25)	27 (21)	11 (37)
Intermediate	285 (29)	49 (38)	6 (20)
High	196 (20)	21 (16)	6 (20)
Very high	149 (15)	19 (15)	3 (10)
Missing	321	42	11
Prior MDS-directed therapy, n (%)	763 (59)	98 (57)	24 (59)
Therapy-related MDS, n (%)	272 (21)	33 (19)	6 (15)
Monosomal karyotype, n (%)	180 (14)	25 (15)	5 (12)
Transplantation-related variable			
Conditioning regimen, n (%)			
Myeloablative	681 (52)	84 (49)	24 (59)
Reduced intensity	491 (38)	74 (43)	17 (41)

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Nonmyeloablative	116 (9)	14 (8)	-
Missing	13	-	-
Donor type, n (%)			
Matched, Related	157 (12)	19 (11)	5 (12)
Matched, Unrelated	730 (56)	107 (62)	26 (63)
Mismatched	256 (20)	33 (19)	7 (17)
Cord Blood	158 (12)	13 (8)	3 (7)
Graft type, n (%)			
Bone marrow	189 (15)	26 (15)	6 (15)
Peripheral blood stem cells	950 (73)	132 (77)	32 (78)
Cord Blood	152 (12)	13 (8)	3 (7)
Other	10 (1)	1 (1)	-
Donor age			
Under 35	790 (61)	110 (64)	28 (70)
35 or older	495 (39)	61 (36)	12 (30)
Missing	16	1	1
Female donor, n (%)	396 (32)	45 (27)	13 (35)
In vivo T cell depletion, n (%)	521 (40)	71 (41)	13 (32)
GVHD prophylaxis, n (%)			
Tacrolimus-based	963 (74)	137 (80)	34 (83)
CSA-based	210 (16)	26 (15)	6 (15)
Other	39 (3)	2 (1)	1 (2)
CD34 selection	31 (2)	1 (1)	-
Ex vivo T-cell depletion	18 (1)	3 (2)	-
Cyclophosphamide-based	18 (1)	1 (1)	-
None reported	22 (2)	2 (1)	-
Year of transplantation			
≤2007	252 (19)	39 (23)	9 (22)
>2007	1,049 (81)	133 (77)	32 (78)

Unless otherwise stated, peripheral blood counts and bone marrow blast counts at time of transplantation

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplemental Table 7. Causes of death by *TERT* rare variant status

	TERT rare variant	
	Present	Absent
	n = 28	n = 902
Cause of death		
GVHD	0 (0)	146 (16)
Non-infectious pulmonary disease	6 (21)	69 (8)
Other malignancy	2 (7)	19 (2)
Organ failure	3 (11)	67 (7)
Primary disease	9 (32)	347 (38)
Infection	5 (18)	139 (15)
Other	3 (11)	115 (13)

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Supplementary Table 8. Multivariable models of transplant outcomes

		Cox regression: overall survival		Competing risks regression: NRM		Competing risks regression: relapse	
	No. of patients	HR (95% CI)	p	HR (95% CI)	p	HR (95% CI)	p
***TERT* rare variant status**							
No rare variant (reference)	1473						
Rare variant	41	1.50 (1.04, 2.17)	0.03	1.75 (1.13, 2.71)	0.01	0.78 (0.42, 1.46)	0.44
***TP53* status**							
No mutation (reference)	1225						
Mutation	289	1.74 (1.49, 2.04)	< 0.001	1.03 (0.82, 1.31)	0.80	1.85 (1.51, 2.28)	< 0.001
IPSS-R Risk Category							
Other (reference)	1343						
Very high	171	1.46 (1.20, 1.78)	< 0.001	1.04 (0.78, 1.38)	0.81	1.41 (1.07, 1.85)	0.01
Donor group							
Matched, Related (reference)	181						
Matched, Unrelated	863	1.07 (0.71, 1.62)	0.73	1.16 (0.63, 2.15)	0.63	1.12 (0.71, 1.77)	0.63
Mismatched	296	1.48 (0.97, 2.25)	0.07	1.69 (0.90, 3.15)	0.10	0.95 (0.59, 1.53)	0.83
Cord Blood	174	1.91 (1.20, 3.03)	0.006	2.19 (1.10, 4.35)	0.03	0.96 (0.56, 1.64)	0.89
RAS-tyrosine kinase pathway mutation							
No mutation (reference)	1321						
Mutation	193	1.35 (1.12, 1.63)	0.002	1.02 (0.77, 1.34)	0.92	1.32 (1.02, 1.71)	0.04
Donor age							
< 35 years old (reference)	928						
35 years or older	568	1.22 (1.05, 1.42)	0.009	1.12 (0.92, 1.37)	0.24	1.00 (0.82, 1.22)	0.99
Missing	18	0.95 (0.50, 1.79)	0.87	0.59 (0.19, 1.85)	0.37	1.06 (0.58, 1.95)	0.86
Recipient age							
10-year increase	1514	1.23 (1.16, 1.30)	< 0.001	1.23 (1.14, 1.33)	< 0.001	0.98 (0.91, 1.04)	0.48
Year of transplant							
2005-2007 (reference)	300						
2008-2014	1214	0.77 (0.49, 1.19)	0.24	0.48 (0.26, 0.90)	0.02	1.86 (1.10, 3.14)	0.02
Karnofsky Performance Score							
90-100 (reference)	817						
10-80	419	1.27 (1.10, 1.48)	0.002	1.23 (1.02, 1.53)	0.03	0.99 (0.80, 1.22)	0.94
Missing	278	1.05 (0.87, 1.27)	0.59	0.95 (0.73, 1.24)	0.72	1.14 (0.90, 1.45)	0.27
HCT-CI							
0 (reference)	258						
1-2	255	1.29 (1.01, 1.65)	0.04	1.15 (0.83, 1.58)	0.40	1.20 (0.88, 1.63)	0.25
3 or above	552	1.47 (1.19, 1.83)	< 0.001	1.32 (0.99, 1.76)	0.06	1.07 (0.81, 1.40)	0.64
Missing	449	1.35 (0.86, 2.11)	0.19	0.78 (0.41, 1.47)	0.44	1.86 (1.12, 3.07)	0.02
Conditioning intensity							
Myeloablative (reference)	789						
Reduced intensity	582	0.90 (0.77, 1.04)	0.15	0.82 (0.67, 1.00)	0.05	1.29 (1.05, 1.58)	0.02
Nonmyeloablative	130	1.04 (0.81,	0.78	0.60 (0.41, 0.90)	0.01	2.19 (1.60, 3.01)	< 0.001

bioRxiv preprint doi: https://doi.org/10.1101/2021.02.11.430624; this version posted February 12, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

		1.33)					
Missing	13	0.99 (0.44, 2.24)	0.98	0.90 (0.28, 2.89)	0.85	2.49 (1.49, 4.16)	< 0.001